DOCUMENT RESUME

ED 095 865 IR 001 045

AUTHOR	Leggate, P.; And Others
TITLE	The BA Previews Project: The Development and Evaluation of a Mechanised S.D.I. Service for Biologists. Final Report. OSTI Report No. 5140.
INSTITUTION	Office for Scientific and Technical Information, London (England).; Oxford Univ. (England). Experimental Information Unit.
REPORT NO	OSTI-5140
PUB DATE	Aug 73
NOTE	288p.
EDRS PRICE	MF-$0.75 HC-$13.80 PLUS POSTAGE
DESCRIPTORS	Biological Sciences; *Biology; *Information Dissemination; Information Retrieval; Interviews; Program Development; Program Evaluation; Questionnaires; Ratios (Mathematics); Relevance (Information Retrieval); Scientific Research; Use Studies
IDENTIFIERS	BA Previews; *Biological Abstracts; Great Britain; SDI; Selective Dissemination of Information

ABSTRACT

During a 2-year period (1970, 1971) SDI (Selective Dissemination of Information) search profiles were written for 353 biologists and other research workers with a need for biological information in academic, industrial, and government research institutions. At the beginning of the experiment a questionnaire and interview survey was made of the information resources then being used by participants, with particular references to their usage of the printed Biological Abstracts (BA). A quantitative description of service performance was provided by the measurement of the following parameters: output sizes, recall ratios, precision ratios, coverage ratios, and novelty ratios. The relevance predictability of the enriched titles which appear on the BA Previews tapes was also studied. The users' reactions to the service were assessed by means of another survey at the end of the experiment. Correlations between quantitative performance measures and the users' assessment of the service, were examined. It was concluded that the results obtained in the project justified a further experiment in which biologists would be asked to make a payment toward the cost of the service, in addition to providing feedback. (Author/WCM)

EXPERIMENTAL INFORMATION UNIT
UNIVERSITY OF OXFORD

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BEST COPY AVAILABLE

THE BA PREVIEWS PROJECT: THE DEVELOPMENT AND EVALUATION OF A MECHANISED S.D.I. SERVICE FOR BIOLOGISTS

P. LEGGATE, B SMITH, J. STOW, M.I. WILLIAMS

August 1973

Experimental Information Unit
7 Keble Road
Oxford OX1 3QL

OSTI REPORT NO. 5140

THE BA PREVIEWS PROJECT:
THE DEVELOPMENT AND EVALUATION OF A MECHANISED S.D.I. SERVICE FOR BIOLOGISTS

A final report to
the Office for Scientific and
Technical Information (Grant S1/24/49)

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
P. LEGGATE
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

by

P. Leggate, B. Smith[1], M.I. Williams[2]
Experimental Information Unit, 7 Keble Road, Oxford

and

J. Stow
Department of Botany, University of Nottingham

AUGUST 1973

IR001045

---

1. Present address: Department of Physiology, University of New England, Armidale, New South Wales, Australia

2. Present address: Science Library (Biological Sciences), The University, Nottingham

Keywords: BA Previews, S.D.I., biologists, evaluation, user studies, recall, precision, coverage.

ABSTRACT

BA Previews, the machine-readable version of Biological Abstracts and BioResearch Index, has been used to provide an experimental S.D.I. service. During a two-year period (1970, 1971) S.D.I. search profiles were written for 353 biologists and other research workers with a need for biological information in academic, industrial and government research institutions. Profiles were also supplied to eight information centres. Three 'liaison scientists' were responsible for recruiting the experimental population, and for interviewing participants in order to identify their information requirements, to construct and amend their search profiles, and to discuss their reactions to the service. At the beginning of the experiment a questionnaire/interview survey was made of the information resources then being used by participants, with particular reference to their usage of the printed Biological Abstracts.

Special studies of the different searchable indexes (enriched titles, CROSS codes, BioSystematic codes) on the BA Previews tapes were made by the liaison scientists in order to increase their knowledge of the data base and hence improve the methods of profile construction.

A quantitative description of service performance was provided by the measurement of the following parameters: output sizes, recall ratios, precision ratios, coverage ratios, and novelty ratios. The relevance predictability of the enriched titles which appear on the BA Previews tapes was also studied. The users' reactions to the service were assessed by means of a questionnaire/ interview survey at the end of the experiment. Correlations between quantitative performance measures and the users assessment of the service, as revealed in questionnaire replies, were examined.

It was concluded that the results obtained in the project justified a further experiment in which biologists would be asked to make a payment toward the cost of the service, in addition to providing feedback.

ACKNOWLEDGEMENTS

The authors wish to thank the following for their interest, help and advice during the project:

1. The members of the advisory committee for the project, under the chairmanship of Dr R.W.J. Keay.
2. Professor E.C.Cocking, grant holder at the University of Nottingham, and Dr N. Tebble and, latterly, Dr A.E. Needham, grant holders at the University of Oxford.
3. The staff of BIOSIS in Philadelphia.
4. The Committee on Biological Information (COBI) who were in part responsible for initiating the project.
5. The following staff of the Experimental Information Unit: R. Firth and J.E. Crowe for programming assistance; E. Evans and R. Jarman for help in data analysis.
6. Dr A.K. Kent (UKCIS) for programming assistance and for making available the UKCIS serial search programs for use in the project.
7. Mrs Diana Simpson and Miss Susan Pusey who typed this report.

Special thanks are due to the biologists who took part in the experiment and who provided the 'feedback' on which this report is largely based.

The project was funded by a grant (S1/24/49) from the Office for Scientific and Technical Information (OSTI).

Notwithstanding the above acknowledgements, the authors take full responsibility for the contents of this report.

Special Note

BA Previews is a trade mark of the BioSciences Information Service of Biological Abstracts (BIOSIS). BioSystematic and CROSS codes are Copyright, 1971 Biological Abstracts.

CONTENTS

Abstract 3

Acknowledgements 4

Contents 5

List of Tables 10

Chapter I: INTRODUCTION 13

1. ORIGINS OF THE EXPERIMENT 13
2. EXPERIMENTAL OBJECTIVES 13
3. OUTLINE OF THE EXPERIMENT 14
4. ADVISORY COMMITTEE FOR THE BA PREVIEWS PROJECT 17

Chapter II: THE DATA BASE 18

A. CHARACTERISTICS OF THE DATA BASE AND SEARCH PROCESS 18

1. GENERAL 18
2. INFORMATION CONTAINED ON THE TAPE FILES 18
3. PROFILE CONSTRUCTION 21
4. COMPUTER OUTPUT 21

B. STUDIES OF CODE AND ENRICHMENT TERM ALLOCATION 23

1. STUDY OF THE EDITORIAL PROCEDURES FOR ADDING KEYWORDS TO AUGMENT TITLES IN BA PREVIEWS 23
    a) Introduction 23
    b) Aims 24
    c) Methodology 24
    d) Results 25
2. STUDY OF BIOSYSTEMATIC CODE USAGE 26
    a) Methodology 26
    b) Results 27

Chapter III: POPULATION OF USERS 30

A. SELECTION OF PARTICIPANTS FOR THE EXPERIMENT 30

1. DEVELOPMENT PHASE OF THE EXPERIMENT (January - March 1970) 30
2. OPERATIONAL PHASE OF THE EXPERIMENT (April 1970 - April 1972) 30
    a) Main recruitment 30
    b) Liaison with users 32
    c) Profile construction 35
    d) Feedback from users 36
    e) Quantitative evaluation phase of the experiment 38

B. CHARACTERISTICS OF PARTICIPANTS IN THE EXPERIMENT WITH SPECIAL REFERENCE TO THEIR LITERATURE SEARCHING HABITS 39

1. GENERAL 39
2. USAGE OF THE LITERATURE 40
    a) Manual literature searching 40
    b) Mechanised information services 42
    c) Non-users of *Biological Abstracts* 44
    d) Users of *Biological Abstracts* 44
3. USERS WHO WITHDREW FROM THE EXPERIMENT 47

Chapter IV: SERVICE PERFORMANCE 49

1. INTRODUCTION 49
2. RELEVANCE DECISIONS 49
3. SAMPLING METHODS 50
4. STATISTICAL METHODS 51
5. PRECISION AND NOVELTY RATIOS 52
    a) Definitions 52
    b) Methods of measuring precision and novelty ratios 52
    c) Mean values for precision and novelty ratios, and output sizes 54
    d) Variables influencing the novelty ratio 55
    e) Additional estimates of precision 56
6. RELEVANCE PREDICTABILITY 57
7. RECALL AND COVERAGE 60
    a) Definitions 60
    b) The reference samples 61
    c) The BV-6 sample (references found by the user) 61

d) Evidence regarding the relative currency of BA Previews provided by the BV-6 samples 62
e) The Alternative Service Samples 63
f) Analysis of Reference Samples 64
g) Estimates of Coverage Ratios 67
h) Estimates of Recall Ratios 70
i) Recall and Coverage for MAJOR and MINOR Value References 71

8. OVERALL RECALL 71
9. RETRIEVAL PERFORMANCE OF PROFILES WRITTEN BY DIFFERENT LIAISON SCIENTISTS 72
10. PERFORMANCE OF PROFILES WRITTEN FOR USERS IN DIFFERENT EMPLOYMENT CATEGORIES 73

Chapter V: USER REACTION TO THE SERVICE 74

1. RESEARCH AIMS 74
2. RESEARCH METHOD 74
3. ADVANTAGES AND DISADVANTAGES OF THE SERVICE (Question 1) 75
4. USERS' OPINIONS OF INDIVIDUAL ASPECTS OF SERVICE PERFORMANCE (Questions 2-6) 78

a) Retrieval Performance of the Profile 79
b) Coverage of BA Previews 80
c) Relevance Prediction 80
d) Profile Construction 81

5. USE MADE OF RELEVANT REFERENCES FOUND BY THE SERVICE (Questions 8 and 9) 82
6. IMPACT OF THE BA PREVIEWS SERVICE ON THE USAGE OF CURRENT PRIMARY, AND SECONDARY JOURNALS (Questions 7 and 10) 83

a) Impact on usage of the printed Biological Abstracts 83
b) Impact on general literature searching habits 85

7. INTEREST IN RETROSPECTIVE SEARCH SERVICE (Question 11) 87
8. INTEREST IN A PERMANENT CURRENT AWARENESS SERVICE (Questions 12 and 13) 88
9. FACTORS INFLUENCING THE USER'S OVERALL JUDGEMENT OF THE VALUE OF THE BA PREVIEWS SERVICE 90

a) Advantages 92
b) Disadvantages 92
c) Use made of the service 93

10. RELATIONSHIP BETWEEN USERS' OPINIONS AND QUANTITATIVE PERFORMANCE MEASURES 93
11. SHARING THE PRINTOUT 94

12. FORM OF COMPUTER PRINTOUT: CARD OUTPUT 94

Chapter VI: INFORMATION CENTRES 95

1. INTRODUCTION 95
2. USER REACTION TO THE SERVICE 96
   - a) Advantages and disadvantages of the service 96
   - b) Service performance 99
   - c) Relevance prediction 99
   - d) Profile construction 99
   - e) Impact of the service on usage of the printed *Biological Abstracts* 100
   - f) Impact of the service on literature scanning by the centre 100
   - g) Retrospective search service 100
   - h) Current awareness service 100
   - i) Usage made of *BA Previews* printout 100
3. SERVICE PERFORMANCE 103
   - a) Precision ratio 103
   - b) Novelty ratio 104

Chapter VII: CONCLUSIONS AND SUMMARY OF RESULTS 105

1. INTRODUCTION 105
   - a) Objectives 105
   - b) The User Population 105
   - c) The Liaison Scientist 105
2. METHODS OF PROFILE CONSTRUCTION 106
3. LITERATURE SEARCHING HABITS OF BIOLOGISTS 108
4. SERVICE PERFORMANCE AND USER REACTION 109
   - a) Output Size, Recall and Precision 109
   - b) Coverage 110
   - c) Novelty Ratios 111
   - d) Overall Recall 112
   - e) Comparative Performance of *BioResearch Index* and *Biological Abstracts* 112
   - f) User Reactions to the Service 112
   - g) Interest in a Retrospective Search Service 114
   - h) The Information Centres 114
5. FUTURE VIABILITY OF A BA PREVIEWS SERVICE 114

BEST COPY AVAILABLE

GLOSSARY OF TERMS USED IN THE EXPERIMENT 117
TIME-TABLE FOR THE EXPERIMENT 119
BIBLIOGRAPHY 123

APPENDICES

Each Appendix contains additional material relating to the Chapter of the same number. For example, Appendix II contains material referred to in Chapter II. A full list of contents appears on the first page of each Appendix.

APPENDIX II 125
APPENDIX III 174
APPENDIX IV 205
APPENDIX V 225
APPENDIX VI 235

LIST OF TABLES

Chapter II

II A Profile construction: usage of data elements 22

Chapter III

III A Pilot study participants 30

III B Establishments, users and profiles by category of employment 33

III C Profiles by specific subject category 34

III D Manual literature searching: secondary journal usage 41

III E Manual literature searching: specialised secondary journals cited by 10 or more users 42

III F Inadequacy of manual literature searching 43

III G Mechanised information services 43

III H Biological Abstracts: reasons for non-usage of Biological Abstracts 44

III I Usage of Biological Abstracts: Indexes 45

III J Reasons for withdrawal from the BA Previews Experiment: replies from 46 deleted users 48

Chapter IV

IV A Sample data for precision and novelty ratios 53

IV B Mean precision and novelty ratios, and output sizes 54

IV C Additional precision ratios for evaluees and non-evaluees 56

IV D Relevance predictabilities (three relevance categories) 58

IV E Relevance predictabilities (two relevance categories) 59

IV F Precision ratios based on title and document relevance decisions 60

IV G Sources in which users found references for the BV-6 sample 62

IV H Dates of appearance in BA Previews of references from the BV-6 samples 63

IV I Recall and coverage ratios - sample sizes 65

IV J Coverage performance of BA Previews 67

IV K % recall, % precision and output size for 80 users 70

IV L Recall and coverage ratios for major and minor value items (BV-6 samples, 80 users) 71

IV M Performance measures for profiles written by the three liaison scientists 72

IV N Performance measures for users in three employment categories 73

Chapter V

V A Advantages of the BA Previews service 76

V B Disadvantages of the BA Previews service 77
V C Retrieval of too many irrelevant references 79
V D Non-retrieval of relevant references 79
V E Journal coverage 80
V F Difficulty in predicting relevance 81
V G Locking up of original articles 82
V H Difficulty in obtaining original articles 83
V I Usage of the printed Biological Abstracts 84
V J Reduction in time spent scanning journals 85
V K Correlation between reduced scanning of journals and concern about time-lag 87
V L Comparative values of retrospective service and current awareness service 88
V M Interest in a permanent current awareness service (Question 12) 89
V N BA Previews service worth £50-£75 per annum (Question 13) 89
V O Significant correlations between replies to Question 12 and 13 and replies to other questions 91
V P Correlations between performance measures and user satisfaction 93

Chapter VI

VI A Information centres: advantages of the BA Previews service 97
VI B Information centres: disadvantages of the BA Previews service 98
VI C Information centres: opinions of service performance 99
VI D Information centres: precision ratios 103
VI E Information centres: novelty ratios 104

TABLES APPEARING IN APPENDICES

[This list does not include the numerous tables which appear in Appendices II.1 and II.2 and relate to the use of CROSS codes and enrichment terms].

Appendix No.

III.1 List of establishments, March 1971 175
III.4 Group and shared profiles 186
III.6 Analysis of comments on feedback cards returned by a sample of 100 users 189
III.8 Evaluees by category of employment 195
III.9 Evaluees by specific subject category 196
III.10 Manual literature searching: Primary journal usage. List of journals cited by 10 or more users 197
IV.6 Distribution of appearance dates in BA Previews of references from coverage samples 218

IV.7 Currency spread in BA/BRI for individual issues of primary journals 220

IV.8 Serial publications not covered by Biological Abstracts or BioResearch Index 221

V.3 Correlations between users' replies to questions 12 and 13 and replies to questions 1-11 232

CHAPTER 1

INTRODUCTION

1. ORIGINS OF THE EXPERIMENT

BA Previews is the machine-readable version of the printed publications, Biological Abstracts and BioResearch Index. Both the printed and the magnetic tape versions are published by the BioSciences Information Service (BIOSIS), Philadelphia, U.S.A.

Three BA Previews tapes are issued each month: two correspond to the semi-monthly issues of Biological Abstracts and one to the monthly issue of BioResearch Index. With the exception of the abstracts, all the information included in the printed publications is present on the tapes and is available for computer searching.

This report describes the development and evaluation of an experimental S.D.I.* service based on the BA Previews data base. The project was financed by grants from the Office for Scientific and Technical Information (OSTI). Prior to the award of these grants, interest in an investigation of the BA Previews service had been expressed by members of the Committee on Biological Information (COBI) under its then Chairman, Dr G.C. Ainsworth. COBI, through the General Secretary of the Institute of Biology, Dr D.J.B. Copp, sought opinion and collaboration from a wide spectrum of biologists. As a result of the encouraging response, it was decided to proceed with the experimental study described in this report. The Committee has since maintained its interest on behalf of the Institute of Biology, the Biological Council and Aslib, and nominated three of its members to serve on the project's advisory committee.

2. EXPERIMENTAL OBJECTIVES

The experiment had the following objectives :

1. To introduce the biological community to a mechanised information service and to test its reaction to this type of information source.
2. To investigate methods of formulating user profiles for searching the BA Previews tapes, making full use of all the searchable indexes offered by the data base.

---

* Selective Dissemination of Information.

3. To measure the performance of the retrieval system in terms of recall, precision and novelty ratios, and similar parameters.
4. To draw conclusions as to the potential demand amongst biologists for a mechanised S.D.I. service and the suitability of BA Previews as the basis for such a service.

In addition, there were two secondary objectives :

5. To collect data on the patterns of literature usage of the biologists participating in the experiment.
6. To identify the level of interest in any possible future service offering retrospective searches of the BA Previews tapes.

Although the project was directed towards the achievement of specific experimental objectives, the service was operational in the sense that it attempted to satisfy the real information needs of research biologists and thereby make a genuine contribution to their research activity.

The experimental service had two features, which were essential to the experiment, but are unlikely to be present in a permanent service :

a) The use of liaison scientists to provide the interface between the user and the service.

b) The provision of a free service.

3. OUTLINE OF THE EXPERIMENT

The staff appointed for the project were three liaison scientists, two located in the Experimental Information Unit at Oxford University and one in the Department of Botany at Nottingham University. They were responsible for the recruitment of participants, the construction and amendment of user profiles and the collection of feedback from users. Each liaison scientist had a first degree in a biological subject; two (B. Smith, J. Stow) had postgraduate research experience and one (M. Williams) had extensive experience in librarianship and information work.

Prior to the recruitment of the main population of users, 55 biologists in the host universities of Oxford and Nottingham took part in a pilot scale study during January-March 1970. The purpose of this study was to provide the liaison scientists with practical experience in techniques of interviewing and profile construction.

With the exception of the host universities, the invitation to participate in the experiment was addressed to the librarians or information officers in academic, industrial and government research institutions. Data collected in a government survey of the employment of biologists was used as a guide to the selection of a population which was reasonably representative with respect to both the employment categories and the subject interests of biologists. Preference was given to establishments within easy travelling distance from Oxford or Nottingham, but users in more distant locations were included if they had expressed a very strong interest in the experiment or if they belonged to under-represented subject fields. Eight information centres took part in the experiment, thereby enabling an assessment to be made of the potential use of the BA Previews service as an information source for retrieving references which could be included in their own abstracting or indexing publications (See Chapter VI).

The personal interview of the liaison scientist with the user was a characteristic of the experiment. With a few exceptions, three interviews were arranged: the initial interview was for the purposes of identifying the user's information needs and usage, and constructing a user profile; the second interview provided an opportunity to optimise the initial profile; and, in the final interview, the user's reactions to the service were assessed by means of a questionnaire.

Recruitment of the user population (353 users; 392 profiles) was completed in December 1970. With the exception of 59 users who withdrew from the experiment, all participants continued to receive three computer printouts per month until May 1972. These printouts corresponded to the three BA Previews tapes received each month from BIOSIS. As a sequel to the present experiment, an OSTI-supported study, incorporating charges to the user, was established at the Department of Botany, University of Nottingham. All participants were given the opportunity to continue their profiles for a further ten months at a cost of £10 and 151 elected to do so. In the second year of this experiment the price will be increased to £25 per annum.

BEST COPY AVAILABLE

In order to provide a better understanding of the data base, studies were undertaken by the liaison scientists concerning the allocation of CROSS and BioSystematic codes, and of "enrichment terms". These studies made a significant contribution to the optimisation of user profiles. In July 1970, before the main recruitment took place, M. Williams spent a week at BIOSIS in Philadelphia in order to acquire detailed information concerning editorial procedures*.

At the first interviews with the users, a questionnaire survey was carried out by the liaison scientists to collect data on the users' existing literature searching habits and, in particular, their attitude to Biological Abstracts. A final questionnaire survey was conducted during the latter stages of the experiment (October-November 1971) to obtain information on the users' reactions to the BA Previews service, its impact on their previous methods of current awareness searching, and their likely interest in any future permanent service offering current awareness and retrospective search facilities.

Throughout the experiment, all participants were asked to return feedback cards, stating the number of relevant references (MAJOR VALUE and MINOR VALUE) retrieved by their profile. These cards also provided an opportunity for users to comment on the service or to suggest amendments to their profile.

However, for the measurement of service performance, much more detailed feedback was required and not all users could be expected to spare the time required for the provision of this additional information. Approximately half the population (151 users) volunteered to take part in this quantitative evaluation phase of the project, which commenced in April 1971. The collection of evaluation data was completed in January 1972. While the evaluation was in progress no amendments to profiles were allowed.

Output size, precision and recall were the parameters used to define retrieval performance. Coverage, novelty ratios and relevance predictability were also measured. Estimates for the performance ratios were based on user judgements of the relevance (MAJOR VALUE, MINOR VALUE, IRRELEVANT) of individual articles.

---

* Report of visit to the United States from 27 July to 6 August, 1970 [Report submitted to OSTI], by Margaret I. Williams. Oxford. 1970. 15p.

The above parameters define the performance, not of the BA Previews data base, but of 'an operational S.D.I. service, based on the BA Previews data base'. The values obtained for recall and precision ratios are therefore dependent on several variables corresponding to the main components which comprise the complete retrieval service. Thus, retrieval performance of a particular profile will be determined by the effectiveness of communication between the user and the liaison scientist, the skill exercised in profile construction, the flexibility of the search logic, and the quality of the indexes provided by the data base. The use of "failure analysis" techniques(8-10) would be necessary to determine the relative impact on retrieval performance of each of these factors. However, the resources devoted to the project did not allow a detailed analysis of this type to be undertaken.

4. ADVISORY COMMITTEE FOR THE BA PREVIEWS PROJECT

An Advisory Committee was set up to advise on the conduct of the project and to assist in liaison with interested organisations. The membership was:

Dr R.W.J. Keay (Chairman) (the Royal Society)

Professor E.C. Cocking (grant-holder on behalf of the University of Nottingham

Mr H.F. Dammers (Shell Research, Sittingbourne, Kent; responsible for processing BA Previews tapes at Shell)

Miss Phyllis I. Edwards (ASLIB, nominated by COBI)

Dr W.M. Henderson (the Agricultural Research Council)

Professor J.B. Jepson (current Chairman of COBI)

Dr A.K. Kent (Technical Director of UKCIS)

Dr A.E. Needham (succeeded Dr N. Tebble as grant-holder on behalf of the University of Oxford)

Dr F.H. Perring (the Institute of Biology, nominated by COBI)

Mr A.G.A. Pickford (the Biological Council, nominated by COBI)

Dr J.E. Smith (member of the Advisory Committee for Scientific and Technical Information (ACSTI))

Dr Anne Warner (nominated by the Biological Council)

CHAPTER II

THE DATA BASE

A. CHARACTERISTICS OF THE DATA BASE AND SEARCH PROCESS

1. GENERAL

BA Previews, the magnetic tape version of Biological Abstracts, consists of 36 tapes/year i.e. 3 tapes/month, two of which correspond to the semi-monthly publication Biological Abstracts and one to the monthly publication BioResearch Index.

In order to search the BA Previews tapes it was decided to use the Serial Search Programs written by the United Kingdom Chemical Information Service (UKCIS) in KDF 9 User Code. These programs had been developed for searching the natural-language records of the Chemical Abstracts Service's data bases and, with a few modifications, could be adapted for use with BA Previews. For the present experiment, it was necessary to transcribe the BA Previews files from the IBM tapes (9 track, 800 b.p.i.) on to KDF 9 tapes: the conversion program was written by UKCIS. Searches were run on the KDF 9 machine at Oxford University.

2. INFORMATION CONTAINED ON THE TAPE FILES

The tape record contains the full bibliographic citation for each document (i.e. author, title, journal, volume number, date and pagination), together with the subject and taxonomic codes allocated by the editors, but not the abstract which appears in the printed Biological Abstracts. Details of the tape format are given in "Guide to BA Previews".(1)

The data elements on the tape which can be searched by the computer are the following: authors, natural-language terms, journal codens, CROSS (subject) codes, and BioSystematic (taxonomic) codes.

(a) Authors

Each author's name is listed, with the surname first, followed by the initials.

(b) Natural-language terms

All the words appearing in the title of the original article are given, followed by 'enrichment terms', i.e. words added by the editors to make the title more meaningful. The BASIC (Biological Abstracts Subjects In Context) index of the printed version of Biological Abstracts is compiled from these enriched titles.

Details of the editorial procedures for enriching titles are given by BIOSIS(1) and the report* of a study of these, which has been carried out during the course of the project, is summarised in this chapter.

(c) Journal Codens

For all references from serial publications, the 5-character ASTM coden is given, followed by the journal abbreviation.

(d) CROSS (Subject) Codes

CROSS (Computer Rearrangement of Subject Specialities) codes are 5-digit codes which are allocated to each reference by the editors to indicate the subject content. Each corresponds to one of the 548 subject headings which are listed in the CROSS Code Index(2).

Difficulties were frequently encountered by the liaison scientists in locating subjects in the CROSS index, which is based on the subject headings of the printed Biological Abstracts and is arranged in code number order. It is not truly hierarchical and there are apparent ambiguities and overlaps between some headings. To facilitate the use of CROSS codes, an alphabetical index was compiled by manual methods. This was incorporated in an "Index to CROSS and BioSystematic Codes"+ which was issued in November 1970 to each establishment participating in the project, in most cases for retention in the library.

CROSS codes are assigned at three levels :

(i) Primary codes: each reference is allocated one primary code only, which corresponds to the printing assignment of the article, i.e. the subject heading under which it appears in the printed Biological Abstracts.

(ii) Secondary codes: there may be one or more of these codes, which indicate other important topics mentioned in the article.

---

* See Appendix II 1

+ An index produced by the Experimental Information Unit (85 pages)

(iii) Tertiary codes: these are used to indicate topics which are mentioned in the article, but are only of marginal interest.

The primary and secondary codes are listed in the CROSS index of the printed Biological Abstracts, but tertiary codes are present on the tape record only.

When the tape records were transcribed to the KDF 9, the following prefixes were added :

MAJ	Primary codes only
SUB	Secondary codes only
CRS	Primary, secondary and tertiary codes

In searching for a specific CROSS code, the use of the prefix CRS in the user profile allowed an unrestricted search to be made for articles allocated this code at any of the three levels of indexing. Alternatively, the use of the prefix MAJ restricted the search to primary codes, whilst use of the prefix SUB restricted the search to secondary codes. As the tertiary codes tend to be assigned very liberally by the BIOSIS indexers, user profiles often included CROSS codes with both MAJ and SUB prefixes to ensure the retrieval of primary and secondary codes but not tertiary codes.

Definitions of the subject areas covered by each code are contained in a BIOSIS in-house publication(3) which was not made available to the project staff. When new codes are introduced, or existing codes are re-defined, a brief definition of their usage is given in the annually-produced "Guide to BA-Previews"(1).

For references where two or more topics appear to be of equal importance, the primary code is assigned in accordance with the instructions contained in another in-house publication(4).

A detailed study of CROSS code usage was carried out by a member of the project staff in an attempt to clarify some of the problems of profile construction which arose during the project. [The results of this study are not included in the present report.]

(e) BioSystematic Codes

BioSystematic codes are 5-digit taxonomic codes which are listed in the BioSystematic Code Index(5). Each of the 750 codes represents a sub-division of the animal or plant kingdom, according to phylum, class or group of related genera. Both living and extinct forms are included. BIOSIS(6) states that "The BioSystematic codes provide access to taxonomic references, since a taxonomic affiliation is made to an appropriate BioSystematic code for every organism reported in a paper, whether or not the organism is mentioned in the title". In the present experiment, the prefix SYS was used to denote a BioSystematic code. An indication is given on the tape record when an organism is mentioned which belongs to a new or rearranged taxonomic group: this was given the prefix NTX.

The results of a detailed study* of BioSystematic code usage are summarised in this chapter.

3. PROFILE CONSTRUCTION

The user profile consisted of a logical arrangement of one or more of the above-mentioned data elements(11). Term weighting was used both to simulate co-ordinate logic and as a device to order the output. In addition to the simple logical operators, and, or and not, use was also made of contextual logic, whereby it was possible to specify that, for retrieval of a reference depending upon linkage of two word terms, the second must occur within a specified number of words of the first.

Table II A shows the usage made of the different data elements by the liaison scientists in profile construction. Nearly all the profiles (96%) included word terms, and most contained co-ordinations of word terms with either CROSS codes (86%) or BioSystematic codes (74%). In the case of CROSS codes, 81% of the profiles included these at the tertiary level of indexing. Primary and secondary CROSS codes were used in 40% and 32% of the profiles, respectively. The average number of search terms was 58 per profile, but it will be noted that there were wide variations among the three liaison scientists. Thus, the average number of search terms/profile used by B. Smith (77.9) was twice the number used by J. Stow (37.8). It was also observed that, in general, Smith's profiles used more complex logic. These differences reflect partly a different philosophy of profile construction but partly the fact that Stow's population included an unusually high proportion of individual university graduate students with rather restricted information requirements.

4. COMPUTER OUTPUT

The computer output was produced in duplicate, the top copy being sent to the user and the second copy incorporated in the user's file, which was

---

* See Appendix II.2

TABLE II A. PROFILE CONSTRUCTION:
USAGE OF DATA ELEMENTS

DATA ELEMENT(S)	No. of profiles in which data element(s) present				
	Liaison Scientist			TOTAL	%
	MIW	BS	JS		
	(108)*	(119)*	(119)*	(346)*	
WORD TERMS ONLY	1	-	4	5	1.5
CROSS CODES ONLY	-	-	4	4	1.2
AUTHOR	37	20	30	87	25.1
CODEN	6	3	9	18	5.2
WORD TERM	103	119	111	333	96.2
CROSS CODE (i.e. MAJ or SUB or CRS)	103	112	92	307	88.7
MAJ	46	78	14	138	39.9
SUB	30	75	9	114	32.9
CRS	88	107	86	281	81.2
BIOSYSTEMATIC CODE	74	103	86	263	76.0
NTX	5	3	2	10	2.9
WORD TERM + CROSS	99	112	86	297	85.8
WORD TERM + BIOSYSTEMATIC	69	103	83	255	73.7
AVERAGE NO. OF SEARCH TERMS/PROFILE	59.8	77.9	37.8	58.4	

* Figures in parenthesis give the total number of profiles which contributed the data.

BEST COPY AVAILABLE

maintained by the liaison scientist.

Accompanying the first output was a profile printout, which was not issued with subsequent runs unless the profile had been amended, i.e. a new profile established. Each output included a search term frequency listing, which showed the number of times each search term in the profile had produced a 'hit'. The number of items on the printout varied from 0-360 references per issue, the average output size being 35 references per issue. [Further details of output size are given in Chapter IV.]

A specimen profile and printout is given in Appendix II.3. A sheet entitled "Explanation of Biological Abstracts Output (1970)"* was distributed to users in September 1970: this interpreted the abbreviations used on the printout.

B. STUDIES OF CODE AND ENRICHMENT TERM ALLOCATION

These studies, which were undertaken to provide a better understanding of the data base, made a significant contribution to the improvement of user profiles. The initial work carried out prior to the main recruitment in September 1970 provided the liaison scientists with guidance concerning the allocation of codes and enrichment terms. Subsequent investigations were made of individual problems which arose in relation to profile construction. The reports+ of the studies of Enrichment terms and Bio-Systematic codes are summarised below.

1. STUDIES OF THE EDITORIAL PROCEDURES FOR ADDING KEYWORDS TO AUGMENT TITLES IN BA PREVIEWS

a) Introduction

As the BA Previews printout contains no abstract, the user must judge the relevance of the source document, in terms of his existing interests, from the title only.

In an attempt to make the title more meaningful, it is the accepted policy of BIOSIS to add keywords. No set thesaurus of terms has been compiled, so that this process involves the selection of natural-language terms by the editors. However, BIOSIS(1) gives some information concerning the categories of keywords which they include in their title enrichment process: "Keywords necessary to clarify or amplify the title are added at the end of the title in the following approximate order :

---

* See Appendix II.4

+ See Appendix II.1 and II.2

- physical descriptor
- organism(s), specific name, common name
- organ system, or tissue used or affected
- geographical location, including geological strata
- drug(s) and affiliation(s) plus the word DRUG(S) linked to the name of a drug by a hyphen
- important chemicals
- specific chemical descriptor (affiliation)
- specific enzyme(s)
- specific instrumentation, apparatus, or methodology (when new, the word NEW is inserted ahead of the term it modifies)
- specific diseases
- purpose of experimentation"

b) Aims

The purpose of this study was to examine some of these categories and determine the extent to which terms are added by the editors. It was to be expected that these added keywords would increase the effectiveness of natural-language searches of BA Previews. From the point of view of profile construction, it was considered important to determine the consistency with which certain categories of terms are added and to ascertain if there was any control of synonyms. On the basis of the findings, it was hoped that some conclusions could be reached as to the degree of confidence which could be placed upon a significant term (e.g. hormone, drug, specific name of organism) being added as an enrichment, if it was not present in the title.

For example, a user who wishes to be alerted to references concerning the drug "Bethanidine" can only hope for 100% retrieval efficiency if he can rely upon "Bethanidine" always being added as an enrichment, even if it is not present in the author title.

c) Methodology

Six categories of enrichment terms were studied :

Physical descriptors: a list of "physical descriptors" was compiled by scanning (a) the KWIC indexes of the enriched titles in Biological Abstracts (the BASIC index) and in BioResearch Index, and (b) a computer-produced frequency list of enrichment terms used in six issues of

Biological Abstracts and BioResearch Index.

Organisms, Geographical location and Drug names: these were investigated by scanning the abstracts in appropriate subject sections of the printed Biological Abstracts for the presence of these categories of terms; their addition as enrichment was then checked in the BASIC index. Checks were also made for the addition of geographical locations to titles in BioResearch Index.

Drug affiliations and Specific instrumentation: these were investigated using output from experimental profiles run against several issues of BA Previews (both Biological Abstracts and BioResearch Index).

d) Results

(i) Physical descriptors

The following terms, which were used to describe publication types, were found added as enrichment: review, abstract, letter, symposium, book, lecture, note, report, thesis, and conference.

(ii) Organism(s): specific name, common name

(a) Common names of organisms are consistently added when not present in the title, but in the abstract.

(b) Specific names of organisms are consistently added when not in the title, but in the abstract.

(c) If both the common name and the specific name are in the abstract, only the latter is added.

(d) If the common name is in the title and the specific name is in the abstract, the latter is added.

(e) If the specific name is in the title and the common name is in the abstract, the latter is not added.

(iii) Geographical locations

(a) In a sample of 250 references, which had the location mentioned in the title, the corresponding country was added in 227 references.

(b) In a sample of 100 references, which had the location mentioned in the abstract, but not in the title, the corresponding country was added in only 24 references. There were some inconsistencies regarding the choice of enrichment terms, e.g. a reference relating to a location in Britain could

BEST COPY AVAILABLE

be enriched with Britain, British Isles, United Kingdom, England, Scotland or Wales.

It was apparent that the enrichment terms were not assigned with sufficient consistency to provide a reliable method of retrieving articles which were concerned with specific geographical locations.

(iv) Drugs

In 95% of the sample of references investigated, specific drugs cited in the abstract were added as enrichment.

(v) Drug affiliations

The affiliations of specific drugs in the title or abstract, e.g. HORMONE-DRUG, appear to be consistently added, if the hormones mentioned are used as drugs in order to observe their effect on a particular system.

(vi) Specific instrumentation

The subject of "electron microscopy" was studied. It was found that this term is not always added to titles of references concerning this topic, even if the term is present in the abstract. Synonyms or related terms such as "electron beam microscopy", "electron micrograph", "fine structure" and "ultrastructure" are also used and these are not necessarily identical to the term appearing in the abstract.

(vii) Importance of enrichment terms in searching BA Previews

In a random sample of 100 references, which included some with non-enriched titles, it was found that an average of 3.1 words or phrases were added to the author title.

In a sample of 150 references, all of which had been retrieved by word terms only, it was found that 31% would not have been retrieved without the presence of the enrichment terms.

2. STUDY OF BIOSYSTEMATIC CODE USAGE

a) Methodology

Studies of the allocation of BioSystematic codes made use of specially-constructed profiles which were run against one or more issues of BA Previews. The output from these experimental profiles included a printout of the codes allocated by BIOSIS to each paper. For some aspects of the study, it was also necessary to scan the corresponding abstracts in the printed Biological Abstracts but, with one exception, reference was not made to the original article.

BEST COPY AVAILABLE

b) Results

(i) Coding omissions

In articles where an organism is mentioned in the enriched title, and/or the appropriate taxonomic CROSS code is allocated, the appropriate BioSystematic code is allocated almost invariably.

(ii) Coding errors

The probability of a BioSystematic code for a particular taxonomic group retrieving an article unrelated to that group is negligible.

(iii) Six studies of BioSystematic code allocation

The allocation of BioSystematic codes was studied in the following six fields of applied biology :

Food Technology: if the type of food involved is mentioned in the enriched title, there is a high probability that the BioSystematic code corresponding to the organism acting as the source of food will be allocated.

Forestry and forest products: the appropriate BioSystematic code was always allocated when specific genera or species were mentioned.

Plant-derived drugs: all but two of 37 articles retrieved were indexed with the appropriate BioSystematic code for the plant from which the drug was derived.

Mycology and antifungal agents: all but five of 40 references allocated the CROSS codes for mycology (CRS36008) and antifungal agents (CRS38508) were also allocated a BioSystematic code for fungi.

Tissue culture: all but two of 112 references allocated the CROSS code for tissue culture (CRS32500) were also allocated the BioSystematic code for the animal or plant from which the cells were originally obtained.

Clinical medicine: all but two of 112 references allocated one of five CROSS codes for clinical medicine were allocated the BioSystematic code for Hominidae (SYS86215).

(iv) Scope of general BioSystematic codes

A study was made of three general codes, i.e. Microorganisms (SYS01000), Chordata (SYS85000) and Vertebrata (SYS85150), to ascertain whether they were assigned only to articles which either describe general principles or

BEST COPY AVAILABLE

fail to mention specific organisms. It was found that 186 out of 243 references did not mention specific organisms; in the remaining 57 references one or more species were mentioned, though general principles were also discussed in half of these.

It was difficult to determine the policy underlying the distinction between Chordata and Vertebrata. However, it was found that no references were allocated both codes.

(v) Overlap between CROSS codes and BioSystematic codes

Microorganisms: analysis of 91 references allocated CRS29500 and/or SYS01000 indicated that the CROSS code was used when a paper dealt with a fundamental activity of microorganisms, whilst the BioSystematic code was used when microorganisms (often unspecified) were relevant to some aspect of the paper, though in many cases this was merely inferred, e.g. infectious diseases.

Paleontology: analysis of printouts of 137 articles retrieved by a profile composed of the three CROSS codes and the three BioSystematic codes for Paleontology showed that, in all but 16 cases, the appropriate BioSystematic code was accompanied by the equivalent CROSS code, e.g. SYS31000 (Paleozoology) was assigned together with CRS63000 (Paleozoology). The general code SYS27000 (Paleobiology) was only assigned to four papers. The general code CRS64500 (Paleobiology) was used for general topics such as stratigraphy and analytical techniques.

(vi) Some recommendations for using BioSystematic codes

In profile construction, the decision whether to use specific names or the appropriate BioSystematic codes will depend upon the number of organisms to be included and the amount of information required about them. The following recommendations are offered:

(a) When broad categories of organisms are required (e.g. Canidae) it is advisable to use BioSystematic codes.

(b) When only one or a few species or genera are required, it is preferable to list specific, generic and common names in the profile, rather than use the appropriate BioSystematic code, which would result in the retrieval of much unwanted material.

BEST COPY AVAILABLE

(c) When only Homo sapiens is of interest, it is preferable to use SYS86215 (Hominidae) instead of listing all the possible synonyms for 'human' (e.g. man, adolescent, girl).

(d) Analysis of a printout of references relating to methods, techniques and theoretical models showed that the general BioSystematic code (e.g. Chordata, Vertebrata) was sometimes used in preference to a specific code (e.g. Hominidae). Thus, either these general codes should be included in such profiles or, preferably, word terms describing specific techniques should be listed.

BEST COPY AVAILABLE

CHAPTER III

POPULATION OF USERS

A. SELECTION OF PARTICIPANTS FOR THE EXPERIMENT

1. DEVELOPMENT PHASE OF THE EXPERIMENT (JANUARY - MARCH 1970)

A pilot group of 55 biologists was recruited from the host universities of Oxford and Nottingham to enable the liaison scientists to acquire practical experience in techniques of interviewing and of constructing user profiles. The two grant-holders invited a number of biologists to seminars, at which the aims of the experiment were described, so that volunteers could be acquired who were willing to act as test users.

TABLE III A. PILOT STUDY PARTICIPANTS

University	Liaison Scientist	Departments	No. of test users
Oxford	M. Williams	Zoology	16
Oxford	B. Smith	Physiology) Biochemistry)	11
Nottingham	J. Stow	Botany) Zoology)	28
TOTAL			55

2. OPERATIONAL PHASE OF THE EXPERIMENT (APRIL 1970 - APRIL 1972)

a) Main recruitment

The main recruitment commenced in September 1970. As users of biological literature include workers in widely differing subject areas (e.g. plant taxonomy, genetics, immunology), it was appreciated that their literature needs would vary considerably. It was also expected that the needs of users in different types of employment would vary. An attempt was therefore made to recruit a population which was as representative as possible of biologists in the United Kingdom.

No up-to-date published statistics were available but, using data obtained in a government survey,* it was possible to obtain estimates of the percentage of workers in different subject fields related to their type of employment. These figures were used as a guide to the selection of a reasonably representative sample of biological research workers in the United Kingdom. The choice of the university departments, and government and industrial research organisations was also influenced by the practical consideration of proximity to Oxford or Nottingham.

As a result of the publicity given to the project by an announcement in OSTI Newsletter (December 1969)(12) and an article in Nature (13 March 1970)(13), a considerable number of enquiries had been received from both research workers and information officers. Wherever possible, their institutions were included, particularly if they belonged to under-represented subject fields.

The majority of the academic population (75%) was recruited from the host universities of Oxford and Nottingham. In the case of the government workers, the main recruitment was from the laboratories of the Ministry of Agriculture, Fisheries and Food and the Agricultural Research Council (ARC). The number of biologists (66) interviewed at ARC establishments was in fact considerably higher than the required figure. However, it was considered valuable to include these, as the ARC intended to make an independent assessment of the value of computer search services to their organisation. Two of the Medical Research Council Units contacted agreed to participate in the experiment. None of the three research associations which were approached wished to join in the experiment.

Eight information centres, which included three Commonwealth Agricultural Bureaux and three Specialised Information Centres, also took part in the experiment. In order to allow for the broad subject coverage of the centres, up to three experimental profiles were allocated to each centre.

In an attempt to include all types of users of biological literature, some research workers in independent research establishments, and four amateur naturalists, were also recruited.

---

* The data collected in this survey was subsequently incorporated in the CSP Biology Manpower Survey. 1972 (Cmnd 4737).

An additional 16 users were recruited by information officers in some of the establishments visited. Most of the responsibility for profile construction and amendment was undertaken by them.

At the completion of the recruitment, a total number of 337 biologists, in 60 establishments, had been interviewed by the liaison scientists and 392 profiles had been written.

Table III B shows the number of types of establishments visited, the users recruited and the profiles written for each category of employment. Table III C shows the number of profiles written in specific subject categories. A list of the 60 establishments visited is given in the Appendix.*

b) Liaison with users

The experimental procedures for liaison with the user, which were developed during the pilot studies, were adopted for the main recruitment stage of the experiment.

With the exception of the host universities, wherever possible, the invitation to participate in the experiment was addressed to the librarian or information officer of the establishment to be visited. An advantage of this approach was that the library usually organised an initial seminar and arranged an interview time-table. Furthermore, it was considered important to involve the library/information department in the provision of mechanised information services such as BA Previews.

On the first visit of the liaison scientist to the institution (e.g. university department, government research laboratory) an introductory talk was usually given to a group of potential participants. The aims of the project, the characteristics of the data base, and the feedback expected from the user, were described, with the aid of a handout+. Individual interviews were then arranged for research workers who were interested in taking part in the project.

The first interview (1-1½ hours) of the liaison scientist with the

---

* See Appendix III.1

+ See Appendix III.2

BEST COPY AVAILABLE

TABLE III B. ESTABLISHMENTS, USERS AND PROFILES BY CATEGORY OF EMPLOYMENT

Employment Category	No. of Establishments	No. of Users	% of Population	No. of Profiles	% of Population
A. Academic	13	143	42 (53)*	164	42 (53)*
B. Independent Research Establishment	3	10	4	13	3
C. Government					
(a) Ministries, NERC etc.	13	49	14 (18)	63	15 (18)
(b) Agricultural Research Council	9	67	19 (8)	73	19 (8)
(c) Medical Research Council	2	6	2 (4)	6	2 (4)
D. Industry	8	48	14 (17)	53	14 (17)
F. Naturalist	4	4	1	4	1
G. Information Centre	8	10	4	16	4
TOTAL	60	337	100 (100)	392	100 (100)

* The percentages in brackets are based on the data obtained in a government survey of the employment of biologists .

TABLE III C. PROFILES BY SPECIFIC SUBJECT CATEGORY

Subject Category	No. of Profiles	%
1. Agricultural Science	23 (48)[*]	5.9 (12.0)[*]
2. Aquatic Biology	10 (8)	2.6 (2.0)
3. Biochemistry	36 (40)	9.2 (10.0)
4. Biometrics	1 (6)	0.3 (1.5)
5. Biophysics	16 (18)	4.1 (4.5)
6. Cell Biology	15 (8)	3.8 (2.0)
7. Ecology and Conservation	24 (16)	6.1 (4.0)
8. Embryology	2 (6)	0.5 (1.5)
9. Endocrinology	12 (6)	3.1 (1.5)
10. Entomology	17 (6)	4.3 (1.5)
11. Experimental Psychology	5 (10)	1.3 (2.5)
12. Food Science	10 (12)	2.6 (3.0)
13. Forestry	5 (6)	1.3 (1.5)
14. Genetics	11 (12)	2.8 (3.0)
15. Immunology	15 (8)	3.8 (2.0)
16. Medicine	5 (40)	1.3 (10.0)
17. Microbiology	18 (24)	4.6 (6.0)
18. Nutrition	1 (2)	0.3 (0.5)
19. Paleontology	5 (4)	1.3 (1.0)
20. Parasitology	9 (4)	2.3 (1.0)
21. Pathology and Forensic Science	9 (10)	2.3 (2.5)
22. Pharmacology and Toxicology	23 (22)	5.9 (5.5)
23. Physiology	35 (28)	8.9 (6.5)
24. Plant Pathology	19 (10)	4.8 (2.5)
25. Plant Physiology	32 (28)	8.2 (7.0)
26. Public Health	1 (2)	0.3 (0.5)
27. Systematic Botany	7 (4)	1.8 (1.0)
28. Systematic Zoology	10 (4)	2.6 (1.0)
29. Veterinary Science	5 (4)	1.3 (1.0)
30. Virology	12 (6)	3.1 (1.5)
Total	392 (400)	100 100

[*] The figures in brackets are the numbers (or percentages) of profiles estimated for each subject category from the data obtained in a government survey of the employment of biologists.

user started with the completion of a questionnaire* designed to ascertain current literature searching practice and attitude to Biological Abstracts. The user was then asked to define his research topic so that a narrative 'Statement of Requirements' could be written.+ A lengthy discussion ensued between the liaison scientist and the user, who was encouraged to give a more detailed description of his research interests, pointing out which topics he wished to be included and any aspects not required. On the basis of the information acquired, the liaison scientist proceeded to compile a user profile.

c) Profile construction

The User Profile consisted of a logical co-ordination of terms or codes from one or more of the following categories:

(i) Words
  (a) appearing in titles of relevant articles
  (b) added by BIOSIS to enrich titles
(ii) CROSS Index (subject) codes
(iii) BioSystematic (taxonomic) codes
(iv) Authors
(v) Journal codens

In computer searching, the terms (or codes) in the profile are matched against the record on the BA Previews tape and a printout of relevant references is produced.

Most of the participants were offered only one BA Previews profile, which attempted in some cases to cover the entire range of their interests, and in others to define one specific area of interest. Some users were allowed to have more than one profile, for one of the following reasons:

(i) The user wanted separate printouts for two or more unrelated topics
(ii) It was impossible to express adequately all the subject requirements of the user within the logical capabilities of one profile.

---

* See Appendix III.3

+ A sample of users' requirements are given in Appendix III.13

BEST COPY AVAILABLE

Conversely, in order to economise in interviewing time and the number of profiles acquired, 35 Group Profiles were written, whereby more than one person attended the interview and contributed to the construction of the profile.

A further 12 Shared Profiles were written, whereby one person only was interviewed, acting on behalf of one or more others. (The sharing of a printout with a research student was not included in this category). Some profiles which were written for one user only were subsequently found to be shared with colleagues. Details of Group and Shared Profiles are tabulated in an Appendix*.

A few weeks after the first interview, the user received a printout, which corresponded either to the semi-monthly issue of Biological Abstracts or to the monthly issue of BioResearch Index.

d) Feedback from users

After the users had received a few initial printouts, a second interview was arranged to discuss the effectiveness of the search and to make any necessary amendments to the profile.

Accompanying each printout was a Feedback card+. This provided an essential means of rapport between the liaison scientist and the user. As well as marking the number of references which were either "Important" or "Relevant" to him, the user was encouraged to write comments on the service and return the card to the liaison scientist. Regular examination of these cards not only provided the liaison scientists with an indication of how the profile was performing but also alerted them to omissions and errors which might otherwise have been overlooked.

To assist users in making a relevance judgement, the following definitions were given:

IMPORTANT: A reference directly related to your research interests, which you will read as soon as possible, i.e. it is essential that you know of it.

---

[*] See Appendix III.4

[+] See Appendix III.5

RELEVANT: A reference which, though not important, is of some interest to you, i.e. you are glad that you were notified of it.

The IMPORTANT category was intended to be interpreted restrictively and the RELEVANT category generously, and the judgement to be independent of whether a reference was already known to the user.

This gave rise to problems for many users, who misunderstood the distinction between these two categories and included the IMPORTANT references in the total number of RELEVANT references. In order to avoid this, it was decided in October 1970 to rename the IMPORTANT and RELEVANT categories as MAJOR VALUE and MINOR VALUE, respectively, and to redefine them as follows:

MAJOR VALUE: Very centrally related to research interests and could make a significant contribution.

MINOR VALUE: Of interest, and happy to have been informed of it, but of no great loss to have missed it.

Subsequent discussions with the users revealed that the renaming of the relevance categories eliminated any misinterpretations.

An analysis was made of the comments on the feedback cards returned by a sample of 100 users; the results are given in an Appendix*. These results can be summarised as follows:

It was found that 88% of the sample made comments on their feedback cards, ranging from 1 to 36 cards/user. From the point of view of the liaison scientists, the most valuable comments were those relating to profile construction. Users appeared to find it more easy to comment upon unwanted subjects than suggest additions to their profile. Twenty two of the 100 users remarked that their printout included unwanted topics, whilst 26 were able to specify terms, codes or journals which should be deleted from the profile. However, only six users suggested additions to their profile. Users also drew attention to spelling errors in their profile and cases of the truncation of word terms causing the retrieval of irrelevant references. After the amendment of their profile, users commented upon its efficacy, expressing either general satisfaction (18), dissatisfaction (4) or indifference (1). General comments were also made upon the efficacy or otherwise of the profile in covering their subject interests.

---

* See Appendix III.6

Some interesting comments were made on the value of the references, 14 users remarking that some MAJOR VALUE references were from journals they did not scan, or had missed in their searching. There were also comments on the non-availability of references. Three users had been unable to obtain references from the NLL. Difficulty was sometimes encountered in making relevance decisions, e.g. distinguishing between MAJOR and MINOR VALUE (4), judging relevance from the title only (2). Two users would have liked a third relevance category to include articles of "fringe interest". Sixteen users made adverse comments on the illegibility of the printout and six could not understand the abbreviations. When a better quality printout was produced, six users commented favourably on this.

e) Quantitative evaluation phase of the experiment

In order to calculate parameters such as precision, recall and coverage (See Chapter IV) it was necessary, over a limited period, to obtain more detailed information than that given on the feedback cards. However, it was appreciated that many users would not be prepared to spare the time to provide further information. Accordingly, on 5 May 1971, a letter* was sent to all users, with a description of how much work would be entailed. They were given three options:

(1) To take part in the quantitative evaluation phase

(2) Not to take part in the quantitative evaluation phase, but to continue to return feedback cards

(3) To withdraw from the experiment.

At the start of the evaluation phase (June 1971) there were 302 users (346 profiles) of whom 151 users (160 profiles) had agreed to be 'evaluees'. Nineteen of these 'evaluee' profiles were Group Profiles (See page 24). Although the evaluees were self-selected on the basis of willingness to co-operate, they proved to be reasonably representative of the subject and employment categories of the total population of users. The employment and subject categories of the evaluees are shown in two tables in the Appendix[*].

---

* See Appendix III.7

Users opting to withdraw from the experiment at this stage brought the total number of profiles deleted since the start of the project to 59. However, eight of these transferred their profile to a colleague.

B. CHARACTERISTICS OF PARTICIPANTS IN THE EXPERIMENT, WITH SPECIAL REFERENCE TO THEIR LITERATURE SEARCHING HABITS

1. GENERAL

Reference has already been made to a questionnaire[+] completed at the first interview. This was designed to ascertain the existing literature searching habits of the biologists participating in the experiment. It was intended for completion by the liaison scientist in discussion with the user. Of the 353 biologists interviewed, 333 provided a completed questionnaire.

No. of users interviewed by liaison scientists up to March 1961 (Questionnaire completed) = 333

No. of users interviewed by liaison scientists up to March 1971 (No questionnaire completed) = 4

No. of users interviewed by information officers up to March 1971 (Questionnaire not completed) = 16

Total no. of users interviewed = 353 (392 profiles)

No. of users who withdrew prior to 30 June 1971 = 59

No. of users who withdrew from experiment, but transferred their profile to a colleague = 8

Total no. of users at 30 June 1971 = 302 (346 profiles)

Participants were asked to indicate on the questionnaire the length (in years) of their research experience. Replies varied from 1 to 45 years but the largest group (32%) were those in the 0-5 year range.

---

[*] See Appendix III.8 and III.9

[+] See Appendix III.3

2. USAGE OF THE LITERATURE

a) Manual literature searching

(i) Primary journal usage

Users were asked to list up to 10 primary journals which they scanned regularly. In all, 698 journals were cited by one or more users.

As might be expected, the most frequently cited journal was *Nature* (159 users)[*], the next three most frequently cited journals were biochemical journals: *Biochemical Journal* (72 users), *Biochimica et Biophysica Acta* (51 users) and *Journal of Biological Chemistry* (47 users). A table in the Appendix[+] lists the primary journals which were cited by 10 or more users.

(ii) Secondary journal usage

A list was given of 10 common secondary journals, and users were asked to indicate which of these they used for current awareness and/or retrospective searching. Their answers are shown in Table III D. It is apparent that *Current Contents* was most widely used for current awareness purposes and *Biological Abstracts* for retrospective searching.

In addition, they were asked to cite any specialised secondary journals (other than those shown in Table III D) which they consulted regularly. One hundred and fifty different titles were cited, of which only 55 were quoted by more than one user. Table III E lists those cited more than 10 times. With the exception of *Bibliography of Reproduction*, they are all Commonwealth Agricultural Bureaux publications.

---

[*] Discussion with users indicated that *Nature* would have been quoted by more, but was not considered to merit inclusion, being classed by them as a journal of general interest, rather than a core journal

[+] See Appendix III.10

BEST COPY AVAILABLE

TABLE III D. MANUAL LITERATURE SEARCHING: SECONDARY JOURNAL USAGE

Title	Current Awareness			Retrospective		
	Prime Tool	Occasional Tool	Total	Prime Tool	Occasional Tool	Total
Biological Abstracts	51	44	(95)*	72	80	(152)*
Chemical Abstracts	31	9	40	44	37	81
Chemical Titles	4	5	9	-	-	-
CBAC	3	3	6	3	2	5
Current Contents	118	39	157	1	2	3
Index Medicus	14	11	25	29	35	64
International Abstracts of Biological Sciences	9	4	13	2	6	8
Excerpta Medica	14	11	25	7	13	20
Science Citation Index	6	4	10	10	22	32
Zoological Record	1	-	1	24	15	39

---

* Compare with replies to specific questions regarding Biological Abstracts (pp. 44,45 and footnote on p. 45).

TABLE III E. MANUAL LITERATURE SEARCHING: SPECIALISED SECONDARY JOURNALS CITED BY 10 OR MORE USERS

Title	No. of Citations
Veterinary Bulletin	26
Horticultural Abstracts	25
Review of Applied Entomology Soils and Fertilisers	18
Helminthological Abstracts	14
Field Crop Abstracts	12
Weed Abstracts	11
Bibliography of Reproduction Nutrition Abstracts and Reviews Review of Plant Pathclogy	10

(iii) General approach to manual literature searching

When asked if they considered their literature searching was adequate, the majority (238 users: 71%) replied "No". [However, it was apparent from comments made at second interviews that some of the 81 users (24%) who had replied "Yes" to this question, had subsequently come to realise that they had not been covering the literature as comprehensively as they originally thought.]

The reasons given by those who said "No" are categorised in Table III F.

b) Mechanised information services

Users were asked if they had previous experience of a mechanised service. Eighty four users replied "Yes" and mentioned the services shown in Table III G. A MEDLARS one-off retrospective search was the service cited by the largest number of users.

BEST COPY AVAILABLE

TABLE III F. INADEQUACY OF MANUAL LITERATURE SEARCHING

Reason	No. of times cited
1. Lack of time	143
2. Diverse interests: too many relevant journals to scan	47
3. Lack of adequate library facilities	46
4. Language barrier: lack of adequate translation facilities	29
5. Lack of a systematic searching policy	22
6. A general feeling of inadequacy	18
7. Inadequate journal coverage (including poor currency, bad format)	16
8. Only a limited number of journals scanned, as a general policy	13
9. No literature searching done: rely entirely upon personal contacts	10
10. Lack of relevant journals, since user is only worker in his field	3

Note: Some users gave more than one reason

TABLE III G. MECHANISED INFORMATION SERVICES

Mechanised Information Service	No. of times cited
(i) MEDLARS (Retrospective)	44
(ii) MEDLARS (Monthly)	5
(iii) ASCA	12
(iv) CT	7
CBAC	17
CAC	3
(v) Ringdoc	13
Agdoc	1
(vi) Drugdoc	1
(vii) Inspec	1

BEST COPY AVAILABLE

c) Non-users of Biological Abstracts

One hundred and fifty five (46%) of those interviewed did not use the printed *Biological Abstracts*. They were asked to give one or more reasons: their replies are shown in Table III H.

TABLE III H. BIOLOGICAL ABSTRACTS: REASONS FOR NON-USAGE OF BIOLOGICAL ABSTRACTS

Reason	No. of times cited
(i) No subscription by the organisation	68
(ii) Inaccessibility	29
(iii) Do not like format of indexes	14
(iv) Do not consider journal coverage adequate for own subject field	9
*(v) Rely on another service	60
(vi) Do not find any need for abstract services	8
(vii) Tend to rely on "Invisible Colleges"	48
(viii) Other	20

* The most commonly cited alternative information service was an internal Library Information Bulletin (11 users), followed by *Current Contents* (8 users) and *Chemical Abstracts* (7 users).

d) Users of Biological Abstracts

The 178 users (54%) who did use the printed *Biological Abstracts* were then asked to state the type of searches they made of it, including the use they made of the indexes, and to give their opinions on certain aspects of the service. Their replies were as follows:

(i) Current Awareness searches

Biological Abstracts was used for current awareness purposes by 105 users (59% of BA users; 32% of the full experimental population of 333 users*). However, only 33 of these 105 users claimed regular usage;

---

* i.e. 68% of the population did not use BA for current awareness purposes (*cf*. p. 84).

the remaining 72 used BA only spasmodically.

(ii) Retrospective searches

One hundred and sixty five participants (92% of BA users; 50% of all users) used Biological Abstracts for retrospective searching, 36 frequently and 129 infrequently*.

(iii) Usage of the indexes

Usage of the indexes to Biological Abstracts is shown in Table III I.

TABLE III I. USAGE OF BIOLOGICAL ABSTRACTS: INDEXES

Index	Current Awareness Searching	Retrospective Searching
Author	44 (42%)*	86 (52%)*
BASIC	51 (49%)+	71 (43%)
CROSS	3 (3%)	8 (5%)
BioSystematic	21 (20%)	25 (15%)
Appropriate subject section	71 (68%)	105 (64%)

* Figures in brackets are percentages of those who use Biological Abstracts, i.e. 105 (Current awareness) and 165 (Retrospective).

The most popular method of searching Biological Abstracts was item-by-item scanning of selected subject sections. This method was adopted by 68% of those who used the journal for current awareness searches and by 64% of those who used it for retrospective searches. Though the majority also made some use of at least one of the four indexes (author, BASIC, BioSystematic, CROSS) there was evidence of substantial under-use of the printed indexes. The BASIC index, which might have been expected to be used extensively for current awareness purposes, was in fact used in only 49% of current awareness searches. Only ten participants (6% of those who used *Biological Abstracts*) claimed to make use of the CROSS index and usage of the BioSystematic index (20% of users) was also lower than expected. In discussion with the participants, the impression was gained that a significant number of biologists either disliked the printed indexes or did not understand how they should be used. As a result they sometimes found it difficult to establish an effective method of searching and failed to locate potentially interesting articles.

---

* Users' replies to this question can be compared with replies to Question 2(a) (ii) regarding secondary journal usage (Table III D). In answer to the latter question only 95 said they used Biological Abstracts for current awareness purposes and 152 that they used it for retrospective searches. It seems that some users, though not considering that Biological Abstracts justified inclusion in their list of secondary journals, when asked a direct question, said that they used it spasmodically or infrequently.

+ Authors note: this figure was wrongly quoted as 45% in the synopsis of this report (OSTI Report No. 5139).

In contrast, nearly all search profiles for *BA Previews* made use of CROSS and BioSystematic codes, which in most cases were linked with other types of index terms, e.g. title words (see Chapter II). Such coordinations are easily accommodated in a computer search, but the equivalent manual search would be extremely laborious.

(iv) Journal coverage

When asked if they considered that the journal coverage of *Biological Abstracts* was adequate for their research interests, their replies were as follows:

Yes	142	(80%)
No	19	(11%)
Do not know No reply	17	(10%)

The interests of those replying "No" covered physical as well as biological sciences, but in addition there were seven core biologists, mainly ecologists, who considered the coverage inadequate.

(v) Quality of the abstracts

Users were next asked if they considered that the abstracts were sufficiently informative: their replies were as follows:

Yes	161	(90%)
No	15	(8%)
Do not know) No reply)	2	(1%)

(vi) Time-lag

Finally, users were asked if the time-lag between publication of an article in a primary journal and its appearance in *Biological Abstracts* rendered the latter inadequate as a current awareness tool. Their replies were as follows:

Yes	93	(52%)
No	81	(46%)
Do not know) No reply)	4	(2%)

3. USERS WHO WITHDREW FROM THE EXPERIMENT

Mention was made on page 27 of 59 users who withdrew from the experiment between February and June, 1971 and whose profiles were deleted (51) or passed on to a colleague (8).

During this period, 10 users withdrew because they were leaving the U.K. for work abroad. The remaining 49 were deleted in two batches, as follows:

(a) February - March 1971: 11 users, who had failed to return feedback cards regularly and did not reply to a written reminder*.

(b) May - June 1971: 38 users, who in response to the letter+ of 5 May (See page 38), opted to withdraw at the start of the quantitative evaluation phase.

Fourteen of the users who withdrew had been recruited in the pilot studies. A total of 33 academic users withdrew, compared with nine

---

[*] See Appendix III.11

[+] See Appendix III.7

government, 14 industrial, and two from independent research establishments. Twenty four of the 30 subject categories were represented, no one particular subject area having an excessive number of dissatisfied users.

In order to ascertain why users withdrew from the experiment, they were sent a questionnaire*, on which they were asked to indicate their reasons. These are shown in Table III J.

TABLE III J. REASONS FOR WITHDRAWAL FROM THE BA PREVIEWS EXPERIMENT: REPLIES FROM 46 DELETED USERS

1. The *BA Previews* service has proved of no value		
(a) Have already seen most of the references	4)	22
(b) The references are too old to be of any use	4)	
(c) The coverage of *BA Previews* is inadequate	4)	
(d) The printout includes too many irrelevant references	10)	
2. Rely on another service		
(a) Abstracting or indexing journal	7)	18
(b) Library/Information service	8)	
(c) Invisible College	2)	
(d) ASCA profile	1)	
3. Analysis of printout and return of feedback is too time-consuming	13	
4. Format of printout is unacceptable	4	
5. Research topic has changed, rendering profile inadequate	4	
6. Leaving the U.K. for work abroad	10	
7. Other	2	

---

* See Appendix III.12

CHAPTER IV

SERVICE PERFORMANCE

1. INTRODUCTION

The following parameters were selected as the most suitable quantitative measures of the performance of the BA Previews S.D.I. service:

(1) Precision: the proportion of references retrieved by the BA Previews service which the user judges to be relevant

(2) Output size: the average number of references retrieved from each 'issue' of the BA Previews data base (there are three 'issues' of BA Previews each month)

(3) Recall: the proportion of relevant references in the data base which are retrieved by the service

(4) Coverage: the proportion of relevant references which are covered by the data base

(5) Overall recall: the proportion of relevant references which are retrieved (irrespective of whether or not they are covered by the data base)

(6) Novelty ratio: the proportion of relevant references retrieved by the service which were *not* previously known to the user

(7) Relevance predictability: the extent to which the document surrogate provided by the service (the enriched title) allows the user to make a correct prediction of the relevance of the actual document

Formal definitions of each of these parameters are given in subsequent sections of this chapter.

One hundred and fifty one users, representing 160 profiles, agreed to take part in the evaluation (*cf*. p. 38), but some evaluees did not return all the required feedback and, in the case of the recall and coverage samples, we did not analyse all the data which was returned. The number of users contributing data to the above performance parameters varies, but is always less than the maximum of 151 users.

2. RELEVANCE DECISIONS

Estimates of performance parameters for each user, with the exception of output size, are derived from the user's relevance decisions on individual references. The majority of relevance decisions were based on a document surrogate rather than on the actual document. In the case of the precision samples, users were asked to decide the relevance of an article from information provided in the computer printout (primarily the enriched title).

It is recognised that the ultimate aim of a retrieval system is to alert the research workers to relevant documents and relevance decisions on the actual documents must usually be considered as definitive. Nevertheless, relevance judgements on the surrogate are considered to be an acceptable basis for estimates of performance ratios. Such judgements correspond to a specific stage in the information process and have real significance to the user. Thus, the user does not refer to all the documents represented in his output but decides, on the basis of the document representation offered by the service, whether further action is justified, e.g. reading the document or noting the reference for future use. Furthermore, routine monitoring of profile performance and the amendment of the search strategy are based on assessments of document surrogates*.

Some relevance decisions based on the full documents were obtained from each user in order to calculate relevance predictability for the surrogate provided in the computer printout from the BA Previews service. These data allowed comparisons to be made of precision ratios derived from the relevance decisions based on the full document and on the enriched title respectively.

Users were asked to classify references as MAJOR VALUE, MINOR VALUE or IRRELEVANT. Definitions of MAJOR and MINOR VALUE categories are given in Chapter III (page 37).

The definitions of performance ratios (recall, precision, coverage and novelty) are expressed in terms of relevant references (relevant = MAJOR VALUE + MINOR VALUE) but similar parameters can be defined for either MAJOR or MINOR VALUE references. Estimates for relevant references are given for all parameters and, in most cases, the corresponding value for MAJOR VALUE items only was also calculated. Values for MINOR VALUE items alone are given only for the recall and coverage ratios.

3. SAMPLING METHODS

Values of performance ratios must be estimated from reference samples. For example, it is not feasible to identify all relevant references in the data base in order to compute a recall ratio. A reasonably simple method must be devised for obtaining a sample of relevant references (the recall sample) which is random with respect to the probability of the reference being retrieved by the profile.

---

* A more detailed discussion of relevance decisions and other aspects of evaluation methodology is given in reference (9).

In order to estimate precision ratios, use can be made of a chronological sample consisting of the output from one or more issues of the data base. Sampling methods for other parameters will be discussed in the appropriate sections.

4. STATISTICAL METHODS(9,14-16)

With the exception of output size all the performance measures are ratios. Mean values for a population of users can therefore be obtained either by averaging ratios or by averaging numbers. In the former method a ratio is computed for each user and these ratios are then averaged. In contrast, a number average is obtained by summing the numerators and denominators separately and then calculating the quotient of the two sums. Thus, for a ratio of the general form $^a/b$ for a population of N users:

$$\text{ratio average} = \frac{1}{N} \times \sum_N \frac{a}{b}\ ; \qquad \text{number average} = \frac{\sum_N a}{\sum_N b}$$

The ratio average is usually preferred as it treats the individual user (or profile) as the sample unit and each user is equally weighted in the calculation of the mean value. If a number average is used, a user who contributes a very large reference sample will have a disproportionate influence on the mean value. If reference samples are of equal size for all users the two methods given identical values.

Unless otherwise indicated, all estimates of mean values reported in this chapter were obtained by averaging ratios. The only major exception is the relevance predictability data given in Table IV D. As the sample sizes for this data varied only slightly (7 - 12 references), the number averages are unlikely to differ significantly from ratio averages.

For some 'group' or 'shared' profiles (cf. Chapter III) performance data were provided by two or more users. Separate ratios were calculated for each user and these were averaged to give a single ratio for the profile which was used in calculating the mean value for the population. The standard errors, which were computed for all ratios averages, define the statistical confidence of an estimate of service performance, i.e. the probability that, within certain error limits, an identical value would be obtained if the experiment were repeated with a different, though equally representative, user population.

The Mann-Whitney U Test[14,15] was used to test for significant differences between two estimates of a performance ratio corresponding to _different_ user populations (e.g. different employment categories). Wilcoxon's matched-pairs signed ranks test[14,15] was used to test for differences between two mean values for the same _user_ population (e.g. recall ratios of MAJOR and MINOR value items). In all cases, a one-tailed test was used and differences were considered significant if the statistical confidence was 95% or greater ($\alpha = 0.05$). In some instances values of α are given in parenthesis to indicate the degree of significance revealed by the test. The above tests were not considered necessary if the difference between two means was greater than _twice the sum_ of the two standard errors. This was taken as sufficient evidence that the difference was significant at the 95% level*.

5. PRECISION AND NOVELTY RATIOS

a) Definitions

$$\text{Precision Ratio} = \frac{\text{number of relevant references retrieved by the service}}{\text{total number of references retrieved by the service}}$$

$$\text{Novelty Ratio} = \frac{\text{number of relevant references not already known to the user}}{\text{total number of relevant refs. retrieved by the service}}$$

Average output sizes (number of references retrieved/issue) were also calculated and are always quoted in conjunction with precision figures. The output size may be at least as significant an indication of user satisfaction as precision. For example, consider two users who are identical with respect to all performance parameters except output size. If the profiles have a precision of 10% but output sizes of 10 and 100 references respectively, it is probable that the former would be considered more satisfactory by the user.

b) Methods of measuring precision and novelty ratios

Estimates of precision and novelty ratios are based on relevance decisions for all references contained in the computer printout from four _BA Previews_ tapes. These tapes corresponded to two issues of Biological Abstracts (52(15), 52(16)) and two issues of BioResearch Index (71(07), 71(08)). For each of these four issues the participants in the evaluation were supplied with two copies of the printout. On one copy they were asked to mark each reference as MAJOR VALUE, MINOR VALUE or IRRELEVANT and, also, to indicate whether they 'already know of the existence' of the MAJOR and MINOR VALUE articles. The latter information was required for the calculation of novelty ratios. The marked copy was then returned to the liaison scientist. The instruction sheet regarding this feedback, which was provided for the guidance of users, is shown in an Appendix*.

---

*For values exhibiting normal distributions, a difference greater than _the sum_ of the two standard errors can be taken as significant at the 95% level. However, performance ratios for information retrieval systems have non-normal distributions. Unless a difference was greater than _twice the sum_ of the two standard errors, the more rigorous non-parametric tests were therefore used.

Precision and novelty ratios were computed for the Biological Abstracts (BA) issues alone, for the BioResearch Index (BRI) issues alone, and for the complete data base (BA/BRI). In computing ratios for the complete data base, data from the BA and BRI issues respectively were weighted in a 2:1 ratio corresponding to their relative frequency of publication. The collection of feedback from two issues each of both BA and BRI was designed to provide equally representative samples of the two parts of the data base. The procedures for calculating precision and novelty ratios for each profile can be illustrated with the figures given in Table IV A for an individual user.

TABLE IV A. SAMPLE DATA FOR PRECISION AND NOVELTY RATIOS

	BA		BRI	
	52(15)	52(16)	71(07)	71(08)
Total no. of references retrieved	33	47	41	51
No. of relevant references (MAJOR or MINOR VALUE)	16	14	8	20
No. of relevant references already known to the user	2	3	4	5

---

* See Appendix IV.1

For BA only

Precision (relevant)[*] $= \dfrac{16 + 14}{33 + 47} =$ 37.5%

Output size $= \dfrac{33 + 47}{2} =$ 40.0 refs/issue

Novelty (relevant) $= \dfrac{(16 - 2) + (14 - 3)}{16 + 14} =$ 83.3%

For BRI only

Precision (relevant) $= \dfrac{8 + 20}{41 + 51} =$ 30.4%

etc.

For BA/BRI

Precision (relevant) $= \dfrac{2(16 + 14) + (8 + 20)}{2(33 + 47) + (41 + 51)} =$ 34.9%

Output size $= \dfrac{2(33 + 47) + (41 + 51)}{6} =$ 42.0 refs/issue

c) Mean Values for precision and novelty ratios, and for output sizes

Table IV B gives mean values for precision and novelty ratios, both for all relevant references and for MAJOR VALUE references, for those evaluees who returned the necessary data. Average output sizes are also given.

TABLE IV B MEAN PRECISION AND NOVELTY RATIOS, AND OUTPUT SIZES

	No. of users[+]	BA	BRI	BA/BRI
Output size	142	34.9±2.8	33.4±3.2	34.4±2.8
No. of relevant items/issue	142	11.2±1.1	10.9±1.2	11.1±1.0
% Precision				
Relevant	142	38.3±2.0	38.7±2.2	38.6±1.9
MAJOR VALUE	142	12.3±1.2	10.9±1.2	12.1±1.1
% Novelty				
Relevant	124	73.4±2.1	84.8±1.8	76.9±1.8
MAJOR VALUE	119	50.8±3.4	74.5±3.2	57.2±3.0

[+] Number of evaluees who returned satisfactory feedback.

---

[*] The terms precision ratio, precision and % precision are treated as synonymous but all numerical values are given as percentages. The same practice is followed for the other performance ratios.

It might be expected that BRI would produce less relevant references than BA but Table IV B shows that the average output size and the mean precision of BA and BRI respectively are almost identical both for all relevant and for MAJOR VALUE references. Predictably, the proportion of articles not previously known to the users was significantly higher for BRI (84.8%) than for BA (73.4%) and the difference was even more marked for MAJOR VALUE references (74.5% for BRI; 50.8% for BA).

One third of all relevant references were judged to be of MAJOR VALUE and the novelty ratio for MAJOR VALUE items (57.2%) was significantly lower than for the relevant (MAJOR + MINOR VALUE) items (76.9%).

d) Variables influencing the novelty ratio

Values for the novelty ratio, which defines the proportion of references found by the service which were not previously known to the user, might appear to be an excellent indicator of user satisfaction, but caution must be exercised if this parameter is to be interpreted in these terms. The novelty ratio of a current awareness service is a composite measure, the value of which is determined by the effectiveness of the service, relative to the user's other current awareness methods, with respect to currency, recall and coverage.

Thus, a fast alerting service such as Chemical Titles or ASCA (Automatic Subject Citation Alert) which may report references before the user sees them in the primary literature, is likely to exhibit a high novelty ratio, irrespective of the recall and coverage performance of the service. On the other hand, the novelty ratio will be determined mainly by recall and coverage performance if the S.D.I. service is less current than a user's 'other methods', as is more likely to be the case for data bases in which there is substantial intellectual input to the indexing and/or abstracting processes. In such cases a high novelty ratio could be taken as an indication either of the retrieval effectiveness of the S.D.I. service, or of the inadequacy of the users' other current awareness methods.

Though no measures of relative currency were made, it is reasonable to assume that the BA Previews service is substantially less current than the user's own current awareness methods. Evidence in support of this assumption is provided by the samples of 'references found by the user's existing current awareness methods' which were obtained for the purpose of computing recall and coverage ratios. This evidence is discussed in a later section of this chapter (page 62). It seems probable, therefore, that the high value (77%) of the novelty ratio for the BA Previews service is determined more by its relative recall and coverage performance than by its relative currency.

It is tentatively suggested that, if the effect of currency can be ignored, the novelty ratio of a service such as BA Previews can be construed as a measure of the effectiveness of the user's conventional current awareness methods. In the present experiment, this would indicate that the user is retrieving only 23% of relevant literature in his subject area by conventional methods. This may in part be due to the fact that users were less conscientious in their manual searching as a result of receiving a BA Previews service.

Replies to the final questionnaire (cf. Chapter V) showed that 92% of users recognised that the service retrieved references which they would not otherwise have found.

e) Additional estimates of precision

The precision ratios given in Table IVB are based on precision samples in which the evaluees identified the individual MAJOR or MINOR VALUE references appearing in their outputs. Additional precision data, for both evaluees and non-evaluees, can be obtained from the feedback cards (Chapter III, page 36) on which the user specified only the number of MAJOR or MINOR value items. Table IVC shows precision ratios for the evaluees, computed from data provided on the feedback cards during two one-month periods (month A and month B) which immediately preceded the collection of the precision sample. Each one-month period corresponds to two issues of BA and one of BRI. Also given in the table are the precision ratios computed from the combined data for both months, for both evaluees and non-evaluees.

TABLE IV C. ADDITIONAL PRECISION RATIOS FOR EVALUEES AND NON-EVALUEES

	Evaluees (140)	Non-evaluees (163)
From feedback cards		
MONTH A	40.2±2.3	-
MONTH B	38.4±2.2	-
MONTHS A and B	39.0±2.2	40.2±2.4
From precision sample*		
	38.6±1.9	-

* cf. page 54

The following conclusions were drawn from these additional precision ratios:

(i) There is no significant difference between the estimates for month A (40.2%) and month B (38.4%). A one month sample is sufficient to provide a valid estimate of precision.

(ii) Estimates (for A and B combined) for evaluees (39.0%) and non-evaluees (40.2%) are almost identical. In relation to precision performance, the evaluees are, therefore, an effectively random sample of the full population.

(iii) For evaluees, the precision ratio provided by the feedback cards (39.0%) does not differ significantly from that obtained from the precision sample (38.6%). Hence, the precision estimates are not influenced by the method of collecting the precision data, i.e. whether a user is asked to give only the number of relevant items, or to identify each reference.

6. RELEVANCE PREDICTABILITY

Relevance predictability is a measure of the extent to which a document surrogate (e.g. a title or an abstract) allows the user to make a correct prediction of the relevance of the actual document. It is expressed in terms of a series of conditional probabilities(17). For example:

$P(T_R \rightarrow D_X)$ = the probability that, given the user has judged a title (T) to be relevant (R), he will judge the document (D) to be irrelevant (X).

The measurement of these probabilities for BA Previews required two relevance judgements on the same item. The user had first to decide the relevance of an article on the basis of the information provided in the BA Previews printout and, subsequently, he had to provide a judgement based on the full document. These dual judgements were obtained for samples of 7 - 12 articles for each user. The samples were a random selection of the English-language articles from serial publications which appeared in the computer printouts for the BA issues 52(15) and 52(16). These printouts formed part of the precision sample (cf. page 54) and evaluees had already provided relevance judgements based on the enriched title which appeared in the printout. The corresponding documents, which were obtained from the National Lending Library, were subsequently distributed to all evaluees.

Relevance judgements on 1049 documents were returned by 133 evaluees[*]. The average interval between judgements on the surrogate and on the actual document was 2 months.

It must be emphasised that relevance predictability data were obtained from reference samples which were not fully representative of the material covered by BA Previews. The samples were restricted to English-language articles from serial publications, and items from BioResearch Index were completely excluded.

Table IV D gives relevance predictabilities calculated from the user relevance decisions. The subscripts A, B and X are used to denote MAJOR VALUE, MINOR VALUE and IRRELEVANT items respectively.

TABLE IV D. RELEVANCE PREDICTABILITIES (THREE RELEVANCE CATEGORIES)[+]

		Sample Size
$P(T_A \rightarrow D_A)$	0.816	135
$P(T_A \rightarrow D_B)$	0.169	'MAJOR VALUE' titles
$P(T_A \rightarrow D_X)$	0.015	
$P(T_B \rightarrow D_A)$	0.185	294
$P(T_B \rightarrow D_B)$	0.682	'MINOR VALUE' titles
$P(T_B \rightarrow D_X)$	0.133	
$P(T_X \rightarrow D_A)$	0.027	620
$P(T_X \rightarrow D_B)$	0.198	'IRRELEVANT' titles
$P(T_X \rightarrow D_X)$	0.775	

[+] All mean values are number averages.

---

[*] The instruction sheet accompanying the documents is shown in Appendix IV 2.

It will be seen from this table that transitions between the MAJOR VALUE and the IRRELEVANT categories are infrequent: only 1.5% (2 items) of MAJOR VALUE titles were subsequently judged to be irrelevant documents, and only 2.7% (17 items) of IRRELEVANT titles were judged as MAJOR VALUE documents. Transitions between MINOR VALUE items and the other categories were more common. However, a clearer overall picture is provided by relevance predictabilities based on a bipartite classification (relevant or irrelevant) as shown in Table IV E. Both number and ratio averages are given in this table and it can be seen that the differences between the two are small.

TABLE IV E. RELEVANCE PREDICTABILITIES
(TWO RELEVANCE CATEGORIES)

	Ratio average	Number average
$P(T_R \rightarrow D_R)$	0.899	0.904
$P(T_R \rightarrow D_X)$	0.101	0.096
$P(T_X \rightarrow D_R)$	0.211	0.225
$P(T_X \rightarrow D_X)$	0.789	0.775
$P(T \rightarrow D)$		0.828*

* the probability of the two decisions being the same

It is evident that the probability that an irrelevant title will correspond to a relevant document (0.211) is significantly greater than the probability that a relevant title will correspond to an irrelevant document (0.101). Furthermore, there are rather more irrelevant titles than relevant ones in a user's output. It appears, therefore, that there is a greater likelihood that the user will miss relevant material because the information provided by the title is insufficient, than that he will waste time in looking up documents which turn out not to be relevant.

The figure in the last row of Table IV E (0.828) defines the probability that the user will make a 'correct' relevance decision based on the title, i.e. it will be the same as his judgement on the full document.

Table IV F shows the two precision ratios for the document sample, computed on the basis of title and document relevance decisions respectively. The precision ratio provided by the full precision sample (_cf_. page 54) is also given.

TABLE IV F. PRECISION RATIOS BASED ON TITLE AND DOCUMENT RELEVANCE DECISIONS

Sample	Relevance Decision	% Precision
Document sample	FULL DOCUMENT	49.1±2.2
	TITLE	40.6±2.4
Precision sample	TITLE	38.6±1.9

Two points of interest emerge:

(i) Precision ratios based on title decisions for the document and precision samples respectively are not significantly different, in spite of the fact that the document sample is limited to English-language articles from serial publications.

(ii) The precision ratio based on document decisions (49.1%) is significantly higher than that for title decisions (40.6%). However, the improved performance suggested by the former ratio is probably illusory from the user's viewpoint. If he is unable to recognise that an article is relevant, he is unlikely to make practical use of it.

7. RECALL AND COVERAGE

a) Definitions

$$\text{Recall ratio} = \frac{\text{number of relevant refs. retrieved by the service}}{\text{number of relevant refs. in the data base}}$$

$$\text{Relative coverage ratio} = \frac{\text{number of relevant refs. in the data base}}{\text{total number of relevant refs. in source X}}$$

where X = a suitable alternative source

The coverage ratio is a relative measure, the value of which will be dependent on the source from which the reference sample was obtained. In contrast, the recall ratio, as defined above, is a characteristic of the service alone and is independent of the source of the recall sample. However, a valid estimate of this measure will be obtained only if the recall sample is equivalent[*] to a random sample of the relevant references in the data base. For example, a search of the printed indexes of BA/BRI would give a non-random ('biased') sample. The manual search would probably use the same index entries which appear in the search profile, and would therefore be likely to retrieve the same references as the profile. A spuriously high recall ratio would result. In almost all circumstances a biased recall sample will give a recall estimate which is significantly higher (rather than lower) than the true value. As will be shown (page 70), at least some of the recall samples obtained in the present experiment were biased.

b) The reference samples

In the BA Previews evaluation, each evaluee was asked to provide two samples of relevant references. Each sample provided estimates of both recall and coverage ratios. The two samples were:

(i) Relevant references found by the user in the course of his normal current awareness searching. These references were listed on a form (BV-6)[+], which was supplied by the liaison scientist, and they will be referred to as the BV-6 sample.

(ii) Relevant references retrieved by an 'alternative service'.

c) The BV-6 sample (references found by the user)

Each evaluee was asked to supply a list of up to 30 relevant (MAJOR or MINOR VALUE) references found in the course of his normal searching of the current literature. The instruction sheet and feedback forms sent to users are shown in an Appendix[+]. Users were also asked to indicate the source from which the reference had been obtained (e.g. primary journal, specified abstract journal). The different sources mentioned by the users and the proportion of references contributed by each source are shown in Table IV G.

---

[*] The sample need be random only with respect to the probability of a reference being retrieved by the service.

[+] Appendix IV.3

TABLE IV G. SOURCES IN WHICH USERS FOUND REFERENCES FOR THE BV-6 SAMPLE

Source	Proportion of references
Primary journals	55%
Current Contents	17%
Abstract journals, reviews	11%
Internal information service (e.g. library bulletin)	9%
Personal communication	6%
Other	2%

d) Evidence regarding the relative currency of BA Previews provided by the BV-6 samples

Our conclusions (page 55) regarding the estimated novelty ratio for BA Previews are dependent on the assumption that BA Previews is, on average, less current than the user's own methods of searching the current literature. The predominance of primary journals and of Current Contents as sources of references found by the user (Table IV G) lends support to this assumption.

More direct evidence is provided by the dates of appearance in the printed issues of BA and BRI of those articles in the BV-6 sample which were found to be covered by the data base. Table IV H summarises the appearance dates for 493 articles which were reported on the BV-6 forms returned by a subset of 42 users. The number of articles appearing in individual months are given in an Appendix*.

---

[*] Appendix IV.6

TABLE IV H. DATES OF APPEARANCE IN BA PREVIEWS OF REFERENCES FROM THE BV-6 SAMPLES

Month of Appearance*	No. of references	% of references
April 1971 or earlier	45	9%
May - Aug 1971	121	25%
Sept - Dec 1971	288	66%
Jan 1972 or later	39	
Total	493	100%

* Based on the printed publication dates given on the printed issues of BA and BRI. The corresponding BA Previews tapes usually arrived in the UK during the same month, but there was a delay of 2 - 6 weeks before the user received the search output.

Most of the BV-6 forms were returned in August 1971 and contained articles found by the user in June, July or August. The data given in Table IV H therefore indicate that, for at least 66% of the BV-6 references, BA Previews is less current than the user's own current awareness methods. For only 9% of the references is it reasonably certain that the computer search would find the reference first.

e) The alternative service samples

All evaluees were provided with an output from one of three alternative services and were asked to mark all MAJOR or MINOR VALUE references on the output, which was then returned to the liaison scientist. The instruction sheets relating to each of the services, which were issued to participants, are shown in an Appendix[+]. (It must be emphasised that the alternative services were being used only to provide recall and coverage samples for the BA Previews service. No attempt was made to optimise search strategies for the alternative services, and our experimental data do not allow comparisons to be drawn between BA Previews and any of these services). The three services were:

(i) A computer search of a 3-month (April - June, 1971) MEDLARS file. This was supplied to 51 users with biomedical interests for whom MEDLARS (Index Medicus) was considered an appropriate alternative data base.

[+] Appendix IV.4

(ii) A computer search of 4 weekly issues of the organic and biochemistry sections of Chemical Abstracts Condensates (CAC), which was supplied to 50 users with biochemical interests. The four issues were those published in February and March 1971 (Vol. 74; issues 5,7,9,11).

(iii) A manual search of Science Citation Index (SCI) for the period January - March 1971. Output from this search was supplied to the remaining 55 participants. Recipients of this service had previously been asked to provide a list of up to 15 important references or authors for use as index entries in the search of the citation index.

Search profiles for the MEDLARS and CAC services were written by the liaison scientists. The SCI search was undertaken by the Institute for Scientific Information (ISI) in Philadelphia.

f) Analysis of reference samples

Reference samples supplied by 80 of the evaluees were analysed. The two reference samples (the BV-6 and the alternative service samples) provided by each of these users were analysed separately. Two recall ratios and two coverage ratios were therefore obtained for each of the 80 users.

Articles from non-serial sources (books, patents, U.S. government reports, conference proceedings, etc) were excluded from both coverage and recall samples. Estimates of both recall and coverage samples were therefore based on relevant references from serial publications only.

Table IV I gives the total number of references (for 80 users) in the different samples which were used in the analysis. The complete sample referred to in Table IV I (row [2]) is the total number of references reported by the users, not all of which were actually analysed.

The instructions given to users (cf. page 61) ensured that no BV-6 sample contained more than 30 references. Furthermore none of the SCI searches yielded more than 30 relevant references. However, the CAC and MEDLARS profiles often produced many more than 30 relevant items. For the purposes of our analysis, the number of items in the CAC and MEDLARS samples was restricted to not more than 35 references by selecting a subset of the complete sample whenever this was necessary. The term experimental sample (Table IV I, [3]) was adopted for the references which were selected for analysis. In the case of the BV-6 and SCI samples, the complete sample and the experimental sample are identical.

TABLE IV I. RECALL AND COVERAGE RATIOS - SAMPLE SIZES

	SOURCE OF SAMPLE				
	BV-6	SCI	CA	MEDLARS	ALL SERVICES
[1] No. of users	80	27	27	26	80
[2] Complete sample	1346	346	868	1548	4108
[3] Experimental sample	1346	346	591	668	2951
[4] Mean sample size (refs/user)	16.8	12.8	21.9	25.7	36.9
[5] No. of non-serial items	63	0	127	0	190
[6] Coverage sample	1283	346	464	668	2761
[7] No. of items covered	950	292	280	482	2004
[8] Recall sample	705	266	184	345	1500
[9] No. of items retrieved	414	119	151	190	874

The Coverage Sample: for each user, the coverage sample consisted of all items in the experimental sample with the exception of non-serial items. Each article in the alternative service samples and in approximately one-half of the BV-6 samples was checked against the printed BA and BRI for the period January 1970 to September 1972 inclusive. Articles from the remaining BV-6 samples were checked for the shorter period January 1970 to February 1972 inclusive. The procedure for determining whether or not an article is covered by BA Previews is detailed in an Appendix*.

A list of the source journals scanned by BIOSIS is published annually[7]. However, not all articles from a journal included in the source list are necessarily selected for inclusion in the data base and, furthermore, articles from a single issue of a primary journal may be distributed over several issues of BA Previews. (Some data on the currency scatter of articles from the same issue of a primary journal is presented in an Appendix+).

---

* Appendix IV.5

+ Appendix IV.7

If an article is not found in the data base during a given period, there is, therefore, no exact criterion for deciding whether or not the article will appear in some later issue of the data base, outside the period during which the data base was searched (we can be certain that an article has not appeared in the data base prior to the period searched, if we search back as far as the publication date of the article, if this is known). However, a rough and ready statistical criterion is provided by the distribution of the appearance dates in BA Previews of those articles which were located in the data base. The distributions of appearance dates for the BV-6 samples and for each of the alternative service samples are given in an Appendix[*]. These data provide strong evidence that the proportion of items appearing outside the period searched is too small to exert any significant influence on the coverage estimates. For example, the 482 'covered' references in the MEDLARS sample were distributed as follows:

1970		1971		1972	
Jan-June	July-Dec	Jan-June	July-Dec	Jan-June	July-Sept
1.9%	1.0%	50.2%	39.6%	6.2%	1.0%

Articles in the BV-6, CAC and SCI samples showed an even smaller proportion of references in the last quarter (July-Sept 1972) of the search period.

The method for computing a coverage ratio for each user, from either of his two samples, can be illustrated as follows: if there are 20 references in the coverage sample and 15 are located in BA Previews then the coverage ratio is 15/20 = 75%.

The Recall Samples: these are subsets of the coverage samples and consist of references known to be covered by the data base, with the following exceptions:

(i) Articles were excluded from the recall sample if they appeared in the BA Previews data base prior to the last profile amendment and hence before the profile was operating at its 'optimum' level.

(ii) Articles in the BV-6 sample (but not the alternative service samples) were excluded if they appeared in an issue of BA Previews from which the user had received a computer printout before he returned the BV-6 form. This ensured that the recall sample did not include any references which the user had actually found in his search output.

---

[*] Appendix IV.6

As shown in Table IV I, out of the 200 references known to be covered by the data base, 504 (25%) were excluded from the recall sample for one or the other of the above reasons.

The *BA Previews* issue in which each article in the recall sample appeared had already been identified during the processing of the coverage sample. Each of these articles was checked against the user's computer printout for that issue in order to determine whether or not the article had been retrieved, and a recall ratio was then calculated. For example, if the recall sample contains 16 references of which 12 appeared in the user's printout, the recall ratio is 12/16 = 75%.

g) Estimates of Coverage Ratios

TABLE IV J. COVERAGE PERFORMANCE OF BA PREVIEWS

	(80 users)	MEDLARS (26 users)	SCI (27 users)	CAC (27 users)
% non-serial items				
[1] alternative service sample	-	0.0	0.0	27.8
[2] BV-6 sample	3.8	-	-	-
Coverage ratio (serial items)				
[3] alternative service	-	72.3±2.8	83.5±4.8	59.2±3.3
[4] BV-6	75.4±2.1	77.4±2.8	72.2±4.6	76.6±3.2
% coverage failures - % from journals not in source list				
[5] alternative service	-	12.6	17.1	38.9
[6] BV-6	15.7	13.4	19.5	14.4

The first two rows of Table IV J give the percentage of articles in the experimental samples which were classified as 'non-serials' and were therefore excluded from both coverage and recall samples.

Coverage ratios for serial items are given in rows [3] and [4] of the table. Mean values are given for each of the three subsets of evaluees corresponding to the three alternative services. For the BV-6 sample, a coverage ratio for the whole population (80 users) is also given. For each subset, coverage ratios were computed both for the alternative service samples and for the BV-6 samples. For example, the column headed 'MEDLARS' gives mean values for the 27 users who received a MEDLARS profile. The figure in row [3] is the % coverage relative to MEDLARS (72.3%); the figure in row [4] is the % coverage relative to the user's own methods of current awareness searching.

Relevant items not covered by BA Previews ('coverage failures') were classified into two groups:

(i) articles from journals not listed in the BA Previews source lists (1970-1972)* and hence not scanned by BIOSIS.

(ii) articles which appeared in journals scanned by BIOSIS but which were not selected for inclusion in the data base.

Rows [5] and [6] of Table IV J show the proportion of coverage failures which were in the former category.

The coverage ratios computed from the BV-6 samples are approximately the same for all three subsets of users. The BA Previews data base covers about 75% of the references found by the user in the course of his normal current awareness searching and, in this respect, users with biomedical or biochemical interests do not differ significantly from core biologists.

The CAC sample gave a lower coverage ratio (59.2%) than the MEDLARS sample (72.3%). This difference is almost entirely accounted for by the number of articles in the CAC sample which were from journals not scanned by BIOSIS. These accounted for 39% of CAC coverage failures as compared to 13% of the MEDLARS and 16% of the BV-6 (all users) coverage failures. It appears that the journal coverage of BA Previews does not extend as far into biochemistry as it does into biomedicine.

However, caution is always necessary in interpreting relative coverage ratios. The coverage ratio of BA Previews relative to CAC is a meaningful measure only if CAC is a reasonable choice of an alternative service for the particular user population (it is assumed that BA Previews is a reasonable choice by virtue of the user's participation in the experiment). Thus, the relatively low value for the coverage ratio would have little significance if CAC was retrieving only 1 relevant item/month compared to the 36 relevant items/month retrieved by BA Previews. In fact, from the complete sample size given in Table IV I, we can compute that CAC is retrieving about 16 articles a month. (This makes no allowance for the fact that we did not attempt to optimise the CAC profile.)

---

* Excluding the 'archival' listings (409 titles in 1971); these are titles which were covered in the past but are now considered unsuitable or are unavailable. Archival titles were classified as 'not covered'.

It must also be noted that we are not in a position to compare the coverage of BA Previews with that of CAC (or of SCI or MEDLARS). Comparisons could be made only if we had also measured the coverage of CAC relative to BA Previews. However, on the basis of the figures given in the previous paragraph it seems possible that many users with biochemical interests will require access to both data bases.

Only a minority of coverage failures were from journals which are not scanned by BIOSIS. For example, only 15.7% of coverage failures from the BV-6 sample were of this type. The majority of coverage failures were articles which had been consciously excluded from the data base either on the grounds of quality or because they were judged to be outside the subject area of the data base.

The journal distribution of the coverage samples was also investigated. The 2004 articles in the coverage samples which were known to be covered by BA Previews were distributed among 648 different primary journals. The coverage sample also contained 141 articles from 86 different journals which do not appear on the BIOSIS source list. A full list of these 86 journals appears in an Appendix*. Three of these journals provided relevant references to three different users, and four provided references to two users. However, the remaining 79 journals contributed relevant material to only one user in each case. There appears to be no case, therefore, for extending the journal coverage of BA Previews.

---

* Appendix IV.8

h) Estimates of Recall Ratios

TABLE IV K % RECALL, % PRECISION AND OUTPUT SIZE FOR 80 USERS

	MEDLARS (26 users)	SCI (27 users)	CAC (27 users)	(80 users)
% Recall				
Alternative service	59.5±4.7	44.0±5.4	82.7±2.9	-
Refs found by the user (BV-6 sample)	54.4±6.1	63.5±5.9	69.1±4.5	62.5±3.2
% Precision	44.9±5.0	39.1±4.4	37.5±3.1	40.1±2.4 (38.6±1.9)[a]
Output Size	39.3±5.8	29.0±5.5	30.1±5.7	32.9±3.3 (34.4±2.8)[a]

a) Mean values for 142 users, *cf* Table IV B

Table IV K shows the recall estimates for the two recall samples supplied by each user. Also given are the % precision and output size for each subset of users. The mean precision and the mean output size for the 80 users for whom recall ratios were computed do not differ significantly from the mean values for the 142 users who supplied precision figures.

We have already pointed out (p. 60) that a valid estimate for the recall ratio, as we have defined it, can be obtained only if the recall sample is 'unbiased' and the estimate is therefore independent of the source of the sample. If this condition is fulfilled then, for a given subset of the population, the BV-6 and the alternative service samples should give values which are identical within the statistical confidence limits of the two estimates.

The recall ratio from the *CAC* sample (82.7%) is significantly higher than that from the BV-6 sample (69.1%) for the same subset of users.

We selected *CAC* as an alternative service because we wished to obtain a relative *coverage* estimate for users with biochemical interests and it had been expected that this sample would give a spuriously high value for the *recall* ratio. Articles are retrieved from *CAC* by searches of author titles and of keywords added by indexers. *CAC* profiles therefore tend to retrieve the same references as those found by searching the enriched titles of BA Previews, and a biased recall sample results.

It was also observed that the recall estimate from the SCI sample (44.0%) is significantly lower than that from the corresponding BV-6 sample (63.5%). It has been argued elsewhere by Leggate(9) that a recall sample can normally only be 'biased' in one direction and this bias will always result in a spuriously high value for the recall ratio. If this view is accepted then we must conclude that the BV-6 sample is biased and gives recall estimates that are higher than the true values. Nevertheless, the remarkably low estimate provided by the SCI sample gives rise to some suspicion as to whether biased sampling may, in some circumstances, result in too low a value for the recall ratio.

If we accept the BV-6 recall ratios as a reasonable basis for comparison between subsets of the population, then it is of interest that profiles for CAC users exhibit higher recall (69.1%) but lower precision (37.5%) than the profiles for MEDLARS users (recall: 54.4%; precision: 44.9%).

i) Recall and Coverage for MAJOR and MINOR value references

TABLE IV L RECALL AND COVERAGE RATIOS FOR MAJOR AND MINOR VALUE ITEMS (BV-6 SAMPLES, 80 USERS)

	% Recall	% Coverage
MAJOR value	66.4±3.8	76.8±2.5
MINOR value	55.9±4.4	74.6±2.7

The figures given in Table IV L illustrate the variations in coverage and recall performance associated with different degrees of relevance. For items in the BV-6 samples the % recall is significantly higher for MAJOR value (66.4%) than for MINOR value items (55.9%), but the coverage ratios for the two categories are effectively the same. A similar picture was presented by the alternative service samples.

8. OVERALL RECALL

Relative overall recall =

$$\frac{\text{number of relevant references retrieved by the service}}{\text{total number of relevant references in source X}}$$

= recall x relative coverage

The recall and coverage ratios correspond to two quite distinct aspects of service performance and the values of each are determined by different sets of variables. Nevertheless, the overall recall ratio, which is the product of recall and coverage ratios, is probably the more meaningful measure from the user's point of view. The user is presumably interested in the proportion of relevant material which is 'missed', irrespective of whether a missed reference represents a recall or a coverage 'failure'.

From the recall and coverage ratios given by the BV-6 samples we computed a value of 47% for the overall recall ratio relative to the user's normal methods. The BA Previews service is retrieving less than half of the relevant literature found by the user's normal searching methods. On the other hand the novelty ratio (77%) indicates that the user was finding only 23% of the relevant material retrieved by the S.D.I. service. It appears that the BA Previews service and the user's own search methods should be regarded as complementary rather than competitive. It is, however, unlikely that the retrieval of all relevant material is necessary for complete user satisfaction, except in the case of information centres and particular types of research workers, e.g. taxonomists.

9. RETRIEVAL PERFORMANCE OF PROFILES WRITTEN BY DIFFERENT LIAISON SCIENTISTS

Table IV M lists the key performance measures for the subsets of profiles written by each liaison scientist.

TABLE IV M PERFORMANCE MEASURES FOR PROFILES WRITTEN BY THE THREE LIAISON SCIENTISTS

	Liaison Scientist		
	JS	BS	MW
Output size (refs/issue)	35.9±7.2 (37)*	32.1±3.2 (62)	37.3±5.5 (43)
% Precision (relevant)	34.3±4.1 (37)	41.0±2.9 (62)	39.8±3.5 (43)
% Recall (relevant, BV-6 sample)	59.3±6.8 (20)	62.2±5.5 (30)	65.1±4.7 (28)
% Novelty (relevant)	79.7±2.7 (26)	78.6±2.4 (53)	75.8±3.1 (42)

* Figures in parentheses indicate the number of profiles which contributed data to each mean value.

The only statistically significant difference among the liaison scientists was in the precision figures for JS and BS ($\alpha = 0.04$). Even this difference is numerically small and we conclude that there are no dramatic differences in the performance of profiles written by different liaison scientists, even though there are substantial differences in the number of terms and the complexity of the search logic (cf p. 21).

10. PERFORMANCE OF PROFILES WRITTEN FOR USERS IN DIFFERENT EMPLOYMENT CATEGORIES

TABLE IV N PERFORMANCE MEASURES FOR USERS IN THREE EMPLOYMENT CATEGORIES

	Employment Category		
	UNIVERSITY	GOVERNMENT	INDUSTRY
Output size (refs/user)	39.8±4.5 (55)	28.7±3.6 (62)	41.0±12.0 (18)
% Precision (relevant)	35.6±2.8 (55)	41.7±2.9 (62)	41.6±7.2 (18)
% Recall (relevant, BV-6 sample)	55.2±4.9 (34)	67.6±5.8 (28)	73.8±7.9 (10)
% Relative Coverage (relevant, BV-6 sample)	80.9±2.7 (34)	71.1±4.0 (29)	68.9±5.5 (10)
% Novelty (relevant)	79.2±2.9 (46)	76.9±2.3 (53)	75.9±4.3 (16)

Statistical tests were made for significant differences between mean values of the performance parameters for university and government users respectively. We found that the BA Previews services supplied to the university population had larger outputs ($\alpha = 0.02$), lower precision ($\alpha = 0.06$), lower recall ($\alpha = 0.05$), and higher coverage ($\alpha = 0.03$). It will be noted that the difference in precision values is not significant at the 95% level. In the case of the other parameters the magnitude of the differences is not large and, furthermore, we can offer no hypothesis to explain these differences. We decided, therefore, that there is insufficient evidence to support the conclusion that government and university research workers represent substantially different populations with respect to S.D.I. services.

A comparison of university and industrial users showed that the latter's profiles exhibited significantly higher recall ($\alpha = 0.03$) and coverage ($\alpha = 0.02$). However, the mean values for % recall and % coverage of industrial users are derived from data for 10 users only and we considered this to be too small a population for any definite conclusions to be drawn.

CHAPTER V

USER REACTION TO THE SERVICE

1. RESEARCH AIMS

The basic purpose of a service such as BA Previews is to satisfy the users' information needs. An essential part of the evaluation was, therefore, the collection of information about users' attitudes to the service. For this purpose a questionnaire* was designed which would try to answer the following questions:

(a) What did the user consider to be the main advantages and disadvantages of the service?

(b) What impact did it have on his existing literature searching practice and, in particular, his usage of the printed Biological Abstracts?

(c) Would he be interested in subscribing to a permanent current awareness service?

(d) Would he be interested in a retrospective search service?

2. RESEARCH METHOD

A pilot study was carried out in August 1971 in order to test a preliminary questionnaire. The latter was sent to 15 selected users who were subsequently interviewed by the liaison scientists. Comments were also invited from the members of the Advisory Committee and a few librarians with experience of user studies.

The final questionnaire was distributed to users in September 1971. A modified version was sent to the eight information centres (See Chapter VI). Users were asked to complete the questionnaire prior to the final interview so that their replies could then be discussed with the liaison scientist. This procedure was designed to prevent the liaison scientist influencing the user's replies, whilst at the same time the final interview provided an opportunity to discuss any difficulties experienced by the user in completing the questionnaire and to clarify ambiguous or apparently inconsistent replies. Written instructions were provided for the guidance of the liaison scientist in these discussions. The minority (10%) of users for whom it was not possible to arrange interviews were asked to return the completed questionnaire by post.

---

* See Appendix V.1

Completed questionnaires were obtained from 280 users. In addition, seven of the eight information centres returned their modified questionnaire.

A computer program was written in order to facilitate the analysis of the questionnaire replies. We were therefore able to undertake a more detailed analysis than would have been feasible by manual methods alone. The program provided counts of individual replies and also of combinations of replies to different questions. It also tested for correlations between different replies by means of the chi-squared test,(15) and for correlations between questionnaire replies and values for performance parameters by computation of Wilcoxon's U-statistic (14,15). Unless otherwise indicated a correlation is described as significant if the significance level (α)* is 0.05 or less ($\geq$ 95% confidence that the correlation is not a random event).

3. ADVANTAGES AND DISADVANTAGES OF THE SERVICE (Question 1)

In order to find out what the users considered to be the main advantages and disadvantages of the BA Previews service, the questionnaire listed seven possible advantages and nine possible disadvantages. Users were instructed to:

".... tick only those which you consider to be really IMPORTANT
to you. Do not tick any which are only of minor significance.
You may tick none, one or several in each list" (Q.1)

Their answers are shown in Tables V A and B.

---

* Two-tailed test.

BEST COPY AVAILABLE

TABLE V A. ADVANTAGES OF THE BA PREVIEWS SERVICE*

Advantage of the service	No. of users ticking this	%
(a) The profile retrieves references of major value which I might not otherwise locate	135	48
(b) The profile retrieves background references which I might not otherwise locate	198	71
(c) The service covers journals which I would not normally see	219	78
(d) The service retrieves references from non-journal publications, e.g. books, symposia, conference proceedings, etc.	103	37
(e) The service is time-saving, i.e. it reduces the time spent searching the literature by normal methods	109	39
(f) The service is convenient, i.e. it obviates the need for copying out references from abstracting and indexing journals for incorporation in my personal index	44	16
(g) The provision of a regular service gives me confidence; my literature searching is no longer dependent upon how and when I can do it myself	88	31

* All but eight users mentioned one or more advantages

BEST COPY AVAILABLE

TABLE V B. DISADVANTAGES OF THE BA PREVIEWS SERVICE*

Disadvantage of the service	No. of users ticking this	%
(j) The profile retrieves too many irrelevant references	120	43
(k) The profile misses too many relevant references	27	10
(l) There is too much delay between publication of an article and its appearance in BA Previews	132	47
(m) The journal coverage is inadequate	26	9
(n) Many relevant references are of little or no practical value:		
(i) they are from journals which are not easily accessible	48	17
(ii) they are from publications which I consider to be of low quality	24	9
(iii) they are from foreign language journals which I cannot read or afford to have translated	60	21
(o) The computer printout is unattractive	78	28
(p) The computer printout is difficult to file	110	39
(q) The coding of the profile is difficult to understand so I am not sure how the profile logic is intended to work, i.e. which single search term or combination of search terms will retrieve a reference	30	11
(r) There is not enough information on the printout:		
(i) no abstracts	75	27
(ii) no author addresses	76	27
(iii) no indication of the language of the original paper	103	37

---

* All but nine users mentioned one or more disadvantages

Advantages: In addition to those listed, three users stated other advantages*.

It appeared that 92% of the population considered that their profile retrieved references which they would not have found by their normal searching methods, as 258 users ticked one or more of the first three advantages (a), (b), (c).

Disadvantages: Two users stated three additional disadvantages*.

4. USERS' OPINIONS OF INDIVIDUAL ASPECTS OF SERVICE PERFORMANCE (Questions 2 - 6)

Questions 2 - 6 were designed to ascertain the users' opinions of individual aspects of service performance.

Questions 2, 3 and 4 referred directly to characteristics which had already been listed as potential disadvantages in question 1 (options j, k, m). However, in questions 2 - 4 the user was being asked whether he considered that the service was deficient with respect to particular performance characteristics (precision, recall, coverage). In question 1 (j, k, m) he had been asked whether each of these characteristics constituted an important disadvantage. The replies to questions 2 - 4 therefore provided a more detailed picture of a user's views than would be offered by question 1 alone.

A comparison of replies to question 1 with those to questions 2 - 4 also provides a test of the consistency of user replies. We assume that a user who mentions a particular deficiency as an important disadvantage must, if he is being consistent, recognise the existence of that deficiency in his replies to questions 2 - 4. For example a user who ticks option 1(j) would be expected to reply 'yes' to question 2.

---

* See Appendix V.2

(a) Retrieval Performance of the Profile

"Do you consider that your profile retrieves too many irrelevant references?" (Question 2)

TABLE V C. RETRIEVAL OF TOO MANY IRRELEVANT REFERENCES

	No. of users	%
Yes	144	51
No	126	45
Don't know	10	4
Total	280	100

Of the 144 users who replied 'yes' to question 2, the majority (118, 82%) had also mentioned the retrieval of irrelevant references as an important disadvantage (Q1(j)). It is a measure of the consistency of user replies that, of the 120 users who mentioned 1(j) as a disadvantage, 118 also replied yes to question 2.

"Do you consider that your profile misses too many relevant references?" (Question 3)

TABLE V D. NON-RETRIEVAL OF RELEVANT REFERENCES

	No. of users	%
Yes	37	13
No	147	52(.5)
Don't know	96	34
Total	280	100

Of the 37 users who replied 'yes' to this question, 25 (68%) had considered 'missed references' (Q1(k)) to be an important disadvantage. Of the 27 users who mentioned option 1(k) as a disadvantage, 25 gave the consistent reply to question 3.

(b) Coverage of BA Previews

"Do you consider that the journal coverage of BA Previews for your research interests is:-" (Question 4)

TABLE V E. JOURNAL COVERAGE

	No. of users	%
Adequate	203	72(.5)
Inadequate	46	16
Don't know	31	11
Total	280	100

Of the 46 users who considered that the journal coverage was inadequate, 24 (52%) had mentioned this as a disadvantage. Two users gave inconsistent replies.

(c) Relevance Prediction

"Do you find it difficult to decide, from the title alone whether an article is relevant to your research interests?" (Question 5)

TABLE V F. DIFFICULTY IN PREDICTING RELEVANCE

	No. of users	%
Frequently	42	15
Occasionally	184	66
No	54	19
Don't know	0	0
Total	280	100

The majority of users (81%) experienced at least occasional difficulty in predicting the relevance of an article, though this was a frequent problem for only 15% of the population.

The users' opinions can be compared with the measurements of relevance predictability which are described in Chapter 4. These showed that there is an 18% probability that a user will not be able to predict correctly the relevance of the document on the basis of the enriched title.

Replies to question 5 were compared with replies to question 1(r)(i): "There is not enough information on the printout: no abstracts". Of the 75 users who considered the absence of abstracts to be an important disadvantage, 20 reported frequent difficulty in predicting the relevance of an article from its title, and 49 reported occasional difficulty. The χ^2 - test showed a very strong correlation ($\alpha = 0.002$) between a positive reply to Q 1(r)(i) and the reply 'frequently' to question 5.

(d) Profile Construction

"The use of a liaison scientist to provide assistance in the construction and amendment of profiles is expensive. Furthermore, some research workers may not wish to delegate this task to someone else but would prefer to construct their own profiles. Would you prefer:

(a) To construct your own profile, after receiving the necessary instruction on the technique of profile construction, and have it checked.

or (b) To have the profile constructed by a liaison scientist, as at present?" (Question 6)

One hundred and fifty eight users (56%) said they would prefer to have their profile constructed by a liaison scientist, the remaining 122 users (44%) would prefer to construct their own profile.

Replies to this question were compared with those to Question 1 (q), "The coding of the profile is difficult to understand". Only 30 users were concerned that they could not understand the coding, and 18 of these said they would be interested in constructing their own profiles.

5. USE MADE OF RELEVANT REFERENCES FOUND BY THE SERVICE (Questions 8 and 9)

"After selecting the relevant references on your computer printout, do you subsequently look up the original articles, provided you have not already seen them?" (Question 8)

TABLE V G. LOOKING UP OF ORIGINAL ARTICLES

	No. of users	%
Always	48	17
Frequently	144	51
Occasionally	85	30
Never	3	1
Total	280	100

It appears that 192 users (69%) make a habit of looking up the original articles and are, therefore, making practical use of the printout.

"Have you experienced difficulty in obtaining the original articles?" (Question 9)

TABLE V H. DIFFICULTY IN OBTAINING ORIGINAL ARTICLES

	No. of users	%
Always	1	-
Frequently	19	7
Occasionally	167	60
Never	92	33
Total	280	100

The answers to this question are not merely a reflection of the available library facilities: users varied in their interpretation of "difficulty". Thus, some users considered that they had "experienced difficulty in obtaining articles" if their library failed to satisfy an inter-library loan request, whilst others considered an article was difficult to obtain if it was not available in their own library, irrespective of whether or not it could be obtained from an outside source.

6. IMPACT OF THE BA PREVIEWS SERVICE ON THE USAGE OF CURRENT PRIMARY AND SECONDARY JOURNALS (Questions 7 and 10)

(a) Impact on usage of the printed Biological Abstracts

"Since you started receiving BA Previews computer printouts, has there been an increase or decrease in your use of the printed Biological Abstracts for current awareness and/or retrospective searching?" (Question 7)

TABLE V I. USAGE OF THE PRINTED BIOLOGICAL ABSTRACTS

	Current Awareness		Retrospective	
	No. of users	%	No. of users	%
Increase	24	9	21	7(.5)
Decrease	52	19	18	6
No change	64	23	116	41
Don't know	5	2	12	4
Never use it	135	48	113	40
Total	280	100	280	100

Participation in the experiment had led to a change in usage of the printed Biological Abstracts for current awareness purposes for 76 users, and for retrospective searching for 39 users. The usage of the printed Biological Abstracts for current awareness and/or retrospective searching increased for 36 users and decreased for 56 users. Thus, participation in the experiment did not result in any major overall change in usage.

Subsequent discussion with users revealed that increased usage for current awareness purposes usually resulted from a favourable reaction to their printout, which led to a desire to consult the abstracts, whilst for retrospective searches it arose from a realisation that Biological Abstracts covered their research interests more adequately than they had hitherto believed. Decreased usage was the result of users finding that they could rely on their printout to retrieve a high proportion of relevant references in their field and that manual scanning of Biological Abstracts was therefore no longer necessary.

It will be noted that the proportion of users (48%) who said they did not use Biological Abstracts for current awareness purposes was lower than that for the first questionnaire (68%) (p. 44). This apparent anomaly is largely accounted for by the fact that 28 of the users who replied that there had been 'no change' and 11 of those who replied that there had been a 'decrease' in usage, had replied 'never use it' in the first questionnaire.

(b) Impact on general literature searching habits

> "Some participants have asked if there will be a permanent BA Previews current awareness service. If there were such a service do you consider that you would be able to reduce the time spent:
>
> (a) Scanning current journals
> (b) Scanning secondary journals." (Question 10)

TABLE V J. REDUCTION IN TIME SPENT SCANNING JOURNALS

	Current journals		Secondary journals	
	No. of users	%	No. of users	%
Considerably	86	31	136	49
Slightly	91	32(.5)	72	26
No	96	34	59	21
Don't know	7	2(.5)	13	5
Total	280	100	280	100

Further analysis showed that only 14% of the population (40 users) believed that a regular BA Previews service would not enable them to reduce the time spent in scanning either the current primary journals or secondary journals.

Replies to this question were compared with those to question 1(e): 'The service is time-saving'. Of 109 users who considered the time-saving feature to be an important advantage, 95 (87%) said they could reduce the time spent scanning current journals.

Replies to question 10 were also correlated with those to question 1 (l) which asked whether the time-lag before the appearance of an article in BA Previews was an important disadvantage. We postulated that a disinclination to reduce the scanning of current journals would be associated with a concern about the time-lag. There should therefore be a strong correlation between a negative reply to question 10 (a) and a positive reply to question 1 (l). We would not, however, expect any similar correlation with the expectation of reduced usage of secondary journals (Q10 (b)) since these would probably not be any more current than BA Previews itself.

The results of the comparison are displayed in Table V K. As expected there is no correlation between replies to question 10 (b) and 1 (l). However, the expected correlation between replies to questions 10 (a) and 1 (l) is weak ($\alpha = 0.07$). We must conclude that the users decision not to reduce his scanning of primary journals must be influenced by other factors than concern about the lack of currency of the S.D.I. service.

TABLE V K. CORRELATION BETWEEN REDUCED SCANNING OF JOURNALS AND CONCERN ABOUT TIME-LAG

Reply to Question 10	% who mentioned time-lag(Q1 (1))*	α = significance level
10 (a), current journals		
users who would not reduce scanning	55	0.07
other users	43	
10 (b), secondary journals		
users who would not reduce scanning	51	0.62
other users	46	

* I.e. the proportion of users giving a particular reply to question 10 who also mentioned time-lag as a disadvantage, e.g. 55% of users who would not reduce their scanning of current journals had mentioned time-lag as a disadvantage.

7. INTEREST IN RETROSPECTIVE SEARCH SERVICE (Question 11)

"If, in addition to the current awareness service, it became possible to operate computer searches of Biological Abstracts for the last 1-5 years, would this retrospective search service be of greater or lesser value to you than the existing current awareness service?" (Question 11)

TABLE V L. COMPARATIVE VALUES OF RETROSPECTIVE SERVICE AND CURRENT AWARENESS SERVICE

Value of retrospective service	No. of users	%
Greater value	117	42
Same value	71	25
Lesser value	71	25
Don't know	21	7(.5)
Total	280	100

There was obviously considerable interest in a service offering retrospective searches as 67% of the users considered they would find this of equal or greater value than the current awareness service.

8. INTEREST IN A PERMANENT CURRENT AWARENESS SERVICE
(Questions 12 and 13)

It was important to discover the degree of interest in a permanent current awareness service, should one become available. Not only was information required concerning the general interest in a regular service, irrespective of cost, but also the reaction to the likely subscription rate for a full cost-recovery service.

"If there were a permanent BA Previews current awareness service would you wish to become a subscriber, assuming that your department/institution would pay the subscription?" (Question 12)

The replies are summarised in Table V M.

TABLE V M. INTEREST IN A PERMANENT CURRENT AWARENESS SERVICE (Question 12)

	No. of users	%
Yes	194	69
No	47	17
Don't know	39	14
Total	280	100

"If a permanent BA Previews current awareness service were established in the U.K., on a full cost-recovery basis, the minimum cost of a profile is likely to be £50-£75 per annum (though improvements in search programs might reduce this substantially during the next few years). Do you consider your present BA Previews profile to be worth £50-£75 per annum?" (Question 13).

TABLE V N. BA PREVIEWS SERVICE WORTH £50-£75 PER ANNUM (Question 13)

	No. of users	%
Yes	113	40
No	109	39
Don't know	58	21
Total	280	100

It was found that of the 194 users who said "yes" to question 12, only 104 (54%) considered that their profile was worth £50-£75 per annum.

An analysis of the employment categories of the users who replied "yes" to question 12 showed that the academic and government workers were more interested in a permanent service (71% and 74% respectively) than their industrial counterparts (50%). However, users who considered the service worth £50-£75 was distributed among the three major employment groups in a significantly different manner. Only 30% of academics considered the service to be worth £50-£75 compared with 49% of government and 44% of industrial biologists. Discussion at interviews showed that replies to question 13 were influenced by the users' views of the likelihood of the necessary funds being made available. Thus, the comparatively low figure for academic workers is explained by the fact that many of these users considered that their department would be unlikely to pay the subscription. On the other hand, a substantial majority of government workers were located at ARC establishments and were hopeful that the ARC would make provision for subscriptions to mechanised services.

9. FACTORS INFLUENCING THE USER'S OVERALL JUDGEMENT OF THE VALUE OF THE BA PREVIEWS SERVICE

Replies to questions 12 and 13 were considered as representing a user's overall opinion of the BA Previews service, taking into account all aspects of its performance. These replies were therefore compared with replies to other questions which were concerned with specific characteristics of the service. The more significant correlations are given in Table V 0. A full list of all the correlations tested appears in an Appendix*.

These comparisons revealed a pattern of factors which apparently influenced the user's wish to subscribe to a permanent service (Question 12) or his opinion that the service was worth £50-£75 (Question 13) but, as expected, no single characteristic had an overwhelming influence on the user's overall judgement of the service. It appeared that the fact that the user mentioned advantages of the service was a better indicator of his reaction than his recognition of specific disadvantages.

---

* See Appendix V.3

BEST COPY AVAILABLE

TABLE V 0. SIGNIFICANT CORRELATIONS BETWEEN REPLIES TO QUESTIONS 12 AND 13 AND REPLIES TO OTHER QUESTIONS

Question	Question 12			Question 13		
	Yes %	No %	α	Yes %	No %	α
Advantages						
Q.1(a) Major value references	60	21	0.0001	62	39	0.0001
Q.1(b) Background references	76	58	0.003	74	68	0.34
Q.1(e) Time-saving	47	20	0.0001	54	29	0.0001
Q.1(f) Convenient	22	1	0.0001	25	10	0.001
Q.1(g) ".... gives me confidence"	40	12	0.0001	47	21	0.0001
Disadvantages						
Q.1(j) Too many irrelevant references	39	52	0.05	32	50	0.003
Q.1(l) Time-lag	45	52	0.30	39	53	0.03
Q.2 [Yes] Too many irrelevant references	46	64	0.01	39	60	0.001
Q.10(a) or (b) [Considerably] Reduce scanning	70	33	0.0001	73	49	0.0001

Notes

(i) The percentages indicate the proportion of those users giving a particular reply to Questions 12 or 13 who also gave the reply indicated in the first column. E.g. the first entry in the column headed "Question 12 yes" indicates that 60% of users replying "yes" to Question 12 had also ticked advantage 1(a).

(ii) α = significance level.

a) Advantages

Strong correlations were found between replies to Questions 12 and 13 and mention of the following advantages :-

(i) Retrieval of MAJOR VALUE references (1(a))
(ii) Time-saving (1(e))
(iii) Convenient (1(f))
(iv) .... the service gives me confidence (1(g))

Users who mentioned "retrieval of background references" (1(b)) as an advantage were more likely to reply "yes" to Question 12 but this advantage did not appear to be related to their judgement of the "worth" of the service (Question 13).

The significance of "time-saving" (1(e)) is further emphasised by replies to Question 10, which showed that 70% of the users who answered "yes" to Question 12, as compared to 33% of those who replied "no", expected that the service would enable a considerable reduction to be made in the time spent scanning either the current journals or secondary journals.

b) Disadvantages

The retrieval of "too many irrelevant references" (1(j);2) was the only disadvantage which showed a significant correlation with the user's overall judgement of the service.

Users who were concerned about the time-lag (1(l)) tended to reply in the negative to Question 13 but the corresponding correlation with Question 12 was not significant.

Further disadvantages of the service, i.e. "misses too many relevant references" (1(k)); inadequate coverage (1(m);L); the provision of references of little practical value (1(n)); the unattractive computer printout (1(o)); and the absence from the printout of abstracts (1(r)-i), author addresses (-ii), or the language of the original article (-iii), did not appear to exert any significant influence on the user's assessment of the value of the service.

c) Use made of the service

It is of interest that users who looked up the original articles (Question 8), and were presumably making the most practical use of the service, were no more likely to reply "yes" to Questions 12 and 13 than those who did not[*].

10. RELATIONSHIP BETWEEN USERS' OPINIONS AND QUANTITATIVE PERFORMANCE MEASURES

Replies to questions 12 and 13 were compared with individual values for the performance measures, i.e. output size, % precision, % recall, % coverage, % novelty. We wished to determine whether user satisfaction with the service, as expressed in positive replies to questions 12 and 13, could be related to 'good' values for a particular performance parameter, e.g. whether users who were interested in a permanent BA Previews service had profiles with above average recall.

The results of correlation tests, using Wilcoxon's U-statistic, are summarised in Table VP which shows, for each parameter, mean values for those who replied yes and no (or don't know) to question 12 and similarly for question 13.

TABLE V P CORRELATIONS BETWEEN PERFORMANCE MEASURES AND USER SATISFACTION

Parameter	Reply to question 12		α	Reply to question 13		α
	Y	N		Y	N	
Output size (refs/issue)	34.0	39.9	(0.79)	36.3	35.4	(0.42)
% Precision	39.8	38.2	(0.56)	41.2	37.8	(0.43)
% Recall[+]	61.5	64.6	*	63.3	62.0	*
% Coverage[+]	69.3	79.5	*	74.5	77.2	*
% Novelty	79.6	69.3	0.001	79.2	73.7	0.006

* Values for α were not computed but are assumed to be non-significant ($\gg 0.05$) on the basis of the mean values given.

[+] BV-6 samples.

---

[*] Details of significance tests appear in Appendix V3.

Only the novelty ratio shows significant correlations with questions 12 ($\alpha = 0.001$) and 13 ($\alpha = 0.006$). This parameter is a rather unsatisfactory measure of service performance because of its composite character (cf p. 55). Nevertheless it is the best predictor of user satisfaction among the parameters which were measured.

11. SHARING THE PRINTOUT

At the commencement of the experiment 47 users (14%) said that they expected to share their printout with colleagues, as a result of which 35 Group Profiles and 12 Shared Profiles were written (Chapter II). It was anticipated that general satisfaction on the part of many users would have led to increased sharing of their printout. Accordingly, users were asked:

"How many people (including yourself) see your printout regularly?" (Q.14)

It was found that 107 users (38%) were sharing their printout with one or more colleagues at the time of the final interviews.

12. FORM OF COMPUTER PRINTOUT: CARD OUTPUT

During the course of the experiment, users showed considerable interest in the possibility of the computer printout being available on cards as opposed to paper. It was considered important to find out what opinion the total population of users had of this card output.

In order to demonstrate the alternative format, one issue of BA Previews (52 (23)) was printed on cards. Accompanying the printout was a short questionnaire* on which users were asked to tick one of three options:

1. I prefer the card output
2. I prefer the paper output
3. I have no preference for cards or paper

There was an overwhelming preference (95%) for the card output.

---

*Appendix V.4

CHAPTER VI

INFORMATION CENTRES

1. INTRODUCTION

In Chapter III reference was made to eight information centres which agreed to take part in the experiment. They were :-

Bee Research Association (Gerrards Cross)

Biodeterioration Centre (Birmingham)

Biomedical Information Centre (Sheffield)

Commonwealth Bureau of Animal Health (Weybridge)

Commonwealth Bureau of Helminthology (St. Albans)

Commonwealth Forestry Bureau (Oxford)

National Documentation Centre for Sport, Physical Education and Recreation (Belfast)

Project FAIR: Specialised Information Centre for Biomedical Engineering (London)

These centres represented a special category of participants in that they served as clearing houses for the collection and subsequent dissemination of information to research workers. They were therefore attempting to satisfy the information needs of others rather than their own individual requirements, as was the case of the ordinary users in the experiment.

Each centre was allowed up to three profiles. In the case of those centres whose interests were confined to biology, an attempt was made to construct a profile which resembled as closely as possible the contents of their information bulletin or abstract journal. The range of interests of some centres extended to subject fields other than biology and therefore BA Previews could not be expected to provide complete coverage.

In common with other participants in the experiment, three interviews were given at each centre in order to construct the initial profiles, to amend these profiles, and to discuss the final questionnaire.

In judging the relevance of the references on their printout most centres found that it was impossible to differentiate between MAJOR VALUE and MINOR VALUE: a reference was either relevant (i.e. selected for incorporation in their system) or irrelevant (i.e. not selected).

2. USER REACTION TO THE SERVICE*

The results of the questionnaire survey of individual users have been described in Chapter V. A modified version of this questionnaire+ was circulated to the eight information centres and, with one exception, these were completed and returned prior to the final interview. From an analysis of the replies given by the seven centres who completed the questionnaire, the following information was obtained :-

a) Advantages and disadvantages of the service (Q.1)

Tables VI A and VI B show the number of centres which considered the advantages and disadvantages listed to be of importance to them. In addition to those listed, the following disadvantages were given :-

(i) "Lack of punctuation in the titles prevents direct incorporation of the reference into our publication."

(ii) "Inadequate name of the source for items from BioResearch Index necessitates consulting the original."

(iii) "It is sometimes difficult to interpret the publication citation, especially for non-periodicals."

---

[*] The authors' comments on questionnaire replies are given in parentheses.

[+] See Appendix VI.1

TABLE VI A. INFORMATION CENTRES: ADVANTAGES OF THE BA PREVIEWS SERVICE

Advantage of the service	No. of centres ticking this*
a) The profile retrieves references which we might not otherwise locate	7
b) The profile retrieves articles from journals which we do not cover in our existing literature searching	7
c) The service retrieves references from non-journal publications, e.g. books, symposia, conference proceedings, etc.	6
d) The service is time-saving, i.e. it obviates the need for manual searching of the hard copy Biological Abstracts	2
e) The service is convenient, i.e. it obviates the need for copying out references from secondary journals, as the printout can be used for processing our bulletin/index	-

---

* Each of the seven centres ticked more than one advantage

BEST COPY AVAILABLE

TABLE VI B. INFORMATION CENTRES: DISADVANTAGES OF THE BA PREVIEWS SERVICE

Disadvantage of the service	No. of centres ticking this*
g) The profile retrieves too many irrelevant references	2
h) The profile misses too many relevant references	3
i) There is too much delay between publication of an article and its appearance in BA Previews	4
j) The journal coverage is inadequate	2
k) The computer printout is unattractive	5
l) The computer printout is difficult to file	1
m) The coding of the profile is difficult to understand	-
n) There is not enough information on the printout:	
(i) no abstracts	1
(ii) no author addresses	3
(iii) no indication of the language of the original paper	4

* Each of the seven centres ticked more than one disadvantage

b) Service performance (Q.2-4)

The centres were asked if they considered that their profile(s) retrieved too many irrelevant references or missed too many relevant references and also if they regarded the journal coverage of BA Previews as adequate. Their answers are shown in Table VI C.

TABLE VI C. INFORMATION CENTRES: OPINIONS OF SERVICE PERFORMANCE

Question	No. of Users		
	Yes	No	Don't know
Profile retrieves too many irrelevant references (Q 2)	3	3	1
Profile misses too many relevant references (Q 3)	4	3	-
	Adequate	Inadequate	Don't know
Journal coverage of BA Previews (Q 4)	3	4	-

[It was to be expected that some information centres would consider that their profile was missing relevant references, as their aim was complete coverage of their interests and they were therefore able to cite references located from other sources, which were relevant to their profile but not retrieved by it. It was also to be expected that the journal coverage would not be comprehensive enough for those centres with multidisciplinary interests.]

c) Relevance prediction (Q.5)

All the centres said that they encountered some difficulty in deciding, from the title alone, whether an article was relevant to their interests, five frequently and two occasionally.

d) Profile construction (Q.6)

Five centres said that they would prefer to construct their own profile, the other two preferring to have it constructed by a liaison scientist.

e) Impact of the service on the usage of the printed Biological Abstracts (Q.7)

Five of the centres never used the printed Biological Abstracts. Of the two centres which did use it, one had decreased the usage for current awareness purposes; the other reported "No change."

[The decreased usage resulted from a realisation that the printed Biological Abstracts need no longer be scanned, an operation which hitherto had taken about an hour per issue. It only took ten minutes to scan the printout, which retrieved relevant references which might have been missed by manual scanning of the printed version].

f) Impact of the service on literature scanning by the centre (Q.8)

Only one centre considered that a regular BA Previews service would enable it to reduce slightly the time spent scanning current journals. Three centres considered they would be able to reduce slightly the time spent scanning the secondary journals. The other three centres stated that no reduction in scanning time would result.

g) Retrospective search service (Q.9)

There was little interest in a retrospective search service. Only one centre said that it would be of the same value to them as the existing current awareness service and none considered it of greater value.

h) Current awareness service (Q.10-11)

Four centres said that, assuming adequate funds were available, they would wish to subscribe to a permanent BA Previews service, but only two of these considered their profile to be worth £50-£75 per annum. The remaining three were uncertain of their interest, though one of them considered their profile to be worth the full cost-recovery rate.

i) Usage made of BA Previews printout (Q.12)

At the final interviews the centres were asked for information on the use they had been making of their printout during the experiment and their general interest in a permanent BA Previews service, should one be established. Their comments were as follows* :-

CENTRE 1: This centre had three profiles, each of which attempted to cover the same subject area as one of three published information bulletins. Some of the relevant references from the printout were incorporated in these three bulletins, which were distributed to 500, 400 and 150 subscribers, respectively. The BA Previews service provided supplementary material, which would not have been located from their other sources, but the poor timeliness of the references prevented the service replacing existing scanning of current journals and of Current Contents. Nevertheless, it had proved sufficiently valuable to justify a subscription.

CENTRE 2: This centre had three profiles designed to cover very specific topics. Some of the relevant references were incorporated in an information bulletin which was distributed free to 250 workers and a few other references were incorporated in the centre's in-house system. The centre was interested in purchasing one or more profiles, but would not be prepared to pay the full rate unless the cost could be shared with another user, i.e. a small group could be generated. They considered that "the viability of a mechanised service such as BA Previews stands or falls on the success of creating groups."

CENTRE 3: This centre had one profile designed to cover the biological aspects of its multidisciplinary interests. Most of the relevant references were incorporated into a published information bulletin, which was distributed to 500 subscribers. It was estimated that the printout retrieved "approximately 300 references/annum which were not previously known." The centre calculated that, assuming a subscription of £60/annum, the cost of each reference would be 20p, which was more than that for Current Contents (7p) but much less than that for their "Cooperating Specialist scheme" (50p). BA Previews could not "replace our other sources, because of the delay factor and must therefore be regarded as an ancillary source, producing at present levels 8%-9% of our total input. The interest in a subscription will depend upon the value put upon this supplementary material."

---

* To preserve anonymity the centres are referred to by numbers; these numbers do not correspond to the order in which the centres are listed on page 95.

CENTRE 4: This centre had one profile designed to cover the entire range of its interests. Some of the relevant references were incorporated in a published abstracting journal, which was distributed to 1,500 subscribers. The centre was interested in subscribing to BA Previews, but was doubtful about the availability of funds. In order to justify purchase of a profile, the centre would have to estimate how many of the relevant references retrieved by the profile would not have been retrieved by existing methods, i.e. scanning of current journals and alerting done by contacts in foreign countries.

CENTRE 5: The profile for this centre attempted to cover the same subject area as the contents of their abstracting journal. Some of the relevant references were incorporated in this journal, which was distributed to 1200-1500 subscribers. They estimated that 10% of the relevant references retrieved by the profile would not have been found by their existing methods; however, some of these would not be selected for inclusion. They were uncertain of their future interest in a BA Previews service and would find it difficult to justify the full cost-recovery rate.

CENTRE 6: This centre had one profile which covered a very broad subject area. Although producing much irrelevant material, the profile retrieved a considerable number of references which would have been missed by their selective manual searching of the hard copy Biological Abstracts. An abstracting journal and an indexing journal were produced by the centre and these were distributed to 2,000 and 500 subscribers, respectively. It was estimated that 5 references/printout were incorporated in the abstracting journal and 15 references/printout were incorporated in the indexing journal. Very few articles of any value were found in BioResearch Index. This centre would be interested in subscribing to a permanent service and would hope to be able to justify the cost.

CENTRE 7: The profile attempted to cover the biological aspects of their range of interests. Only a few of the relevant references retrieved by the printout were incorporated in their abstracting journal, which was distributed to 1,190 subscribers. They said "We have treated the project by and large as an experimental, not a practical, exercise. We have used it to borrow one or two articles of peripheral interest that we might not otherwise have seen." They therefore showed little interest in a permanent service, as they preferred to rely on their existing methods.

CENTRE 8: No questionnaire returned.

3. SERVICE PERFORMANCE

a) Precision ratio

The precision ratio (MAJOR + MINOR VALUE items) was estimated for each of the eight information centres from relevance judgements given on feedback cards. Both the average output size and the precision figures were based on data from six issues of BA Previews, four corresponding to Biological Abstracts and two to BioResearch Index. The results are shown in Table VI D.

TABLE VI D. INFORMATION CENTRES: PRECISION RATIOS

Information Centre	Output Size*	Precision Ratio %
Centre 1	84	33+
Centre 2	54	21
Centre 3	73	24
Centre 4	36	45
Centre 5	28	87
Centre 6	76	14
Centre 7	148	66
Centre 8	14	18
Average	86	39

* Average number of references/issue

+ This figure is the average for 3 profiles

b) Novelty ratio

Novelty ratios were computed for the five information centres who took part in the evaluation phase of the experiment and provided data on the proportion of relevant references retrieved by their profile which they had not already seen. The results are given in Table VI E.

TABLE VI E. INFORMATION CENTRES: NOVELTY RATIOS

Information Centre	Novelty Ratio (%)		
	BA	BRI	BA/BRI*
Centre 1	69+	84+	77+
Centre 3	35	81	47
Centre 5	33	82	41
Centre 6	62	47	54
Centre 7	16	17	16
Mean	45	63	49

* From 4 issues of BA and 2 of BRI (cf. Chapter IV)

+ Mean value for two profiles

As in the case of individual users, BioResearch Index (BRI) supplied a much higher proportion of references not already known to the centre than did Biological Abstracts (BA). Novelty ratios showed wide variations among the individual centres but, with one exception, were substantially lower than the average value (77%) for individual users, as would be expected.

CHAPTER VII

CONCLUSIONS AND SUMMARY OF RESULTS

1. INTRODUCTION

a) Objectives

It will be apparent that the project achieved its first objective of developing an experimental SDI service based on the computer searching of BA Previews, and thereby introducing a large population of biologists to a mechanised service, in most cases, for the first time. The other objectives defined in Chapter I were: to investigate methods of formulating user profiles for searching the BA Previews data base; to assess user reactions to the BA Previews service; to obtain quantitative measures of performance; and to draw conclusions as to the demand among biologists for a mechanised S.D.I. service and the suitability of BA Previews as the basis for such a service. Secondary objectives were the collection of information on literature usage by biologists, and an assessment of the interest in a retrospective search service.

Before summarising the more important results obtained and the conclusions which can be drawn in respect to these objectives, some general observations are necessary on: (a) the choice of the user population, and (b) the role of the liaison scientist in the project.

b) The User Population

The results obtained in relation to each of the objectives listed above are applicable to the biological community as a whole only to the extent to which participants in the project formed a representative experimental population. Figures obtained from a survey of the employment type and subject interests of biologists were used as a guide in the selection of participants. This method of selection ensured that the final population of 337 users was reasonably representative of the major employment categories and the major subject groupings of biologists. It is assumed that the 151 users who participated in the evaluation were a random subset of the full population.

c) The Liaison Scientist

The use of 'liaison scientists' to maintain contact between the user and the service and, in particular, the emphasis on personal interviews with participants, was characteristic of the BA Previews Project and also of other projects undertaken by the Experimental Information Unit[10, 18].

The amount of effort which was devoted to interviewing users, to the definition of their information needs, and to the formulation and maintenance of profiles, was almost certainly greater than would be feasible in a permanent cost-recovery service. The intensive effort devoted to each user is very desirable during the development and evaluation of a new, experimental service, and was probably a significant factor in ensuring the continuance of user interest and co-operation throughout the two years of the project. It was not possible to assess whether the reduced level of assistance, which would be offered to users by the supplier of a permanent cost-recovery service, would result in a significant reduction in the quality of the search profiles.

It seems probable that, in any future permanent service, either the user would have to take greater responsibility for the construction of his own profile, or the liaison function would have to be provided by a local information officer or librarian. In either case the provision of effective training in methods of profile construction for either the user or the information officer would be essential.

2. METHODS OF PROFILE CONSTRUCTION

The BA Previews data base contains both natural language (the enriched titles) and controlled language (CROSS code and BioSystematic) indexes. Much of the first year of the project was devoted to gaining experience in the use of these indexes in the construction of user profiles. Special studies were undertaken by the liaison scientists of the BIOSIS indexing philosophy in selected subject areas as it applied to enrichment terms, BioSystematic codes, and CROSS codes.

Many of the profiles written in the early stages of the project underwent major amendments as our knowledge of the data base and our methods of profile construction improved. The final versions of the search profiles made extensive use of all three indexes. All but five profiles contained word terms, and the CROSS and BioSystematic codes were used in 89% and 76% of profiles respectively. A comparison of profiles written by the three liaison scientists showed dramatic variations both in the number of terms and in the complexity of the profile logic, but these variations were not reflected in any substantial differences in the retrieval performance of the profiles.

Enrichment Terms: these are added by BIOSIS indexers in order to make the titles more meaningful, and their presence undoubtedly enhances the effectiveness of word searches. A small scale study of 150 references, all of which had been retrieved by word terms alone, showed that 31% would not have been retrieved without the presence of the enrichment terms. However, it must always be kept in mind that the enrichment terms are a natural-language index and that allowance must be made for the use of alternative grammatical forms, synonyms and related terms. A special study was made of the potential value of enrichments for specifying geographical locations. This type of enrichment could be especially useful for users interested, for example, in environmental biology. However, it was found that the enrichment terms were not assigned with sufficient consistency to provide a reliable method of retrieving articles which were concerned with specific geographical locations.

BioSystematic Codes: a study of BIOSIS indexing practice for selected BioSystematic codes showed that these codes are applied with a high degree of consistency. This is reflected by the fact that taxonomists were among the most satisfied users of the service. The code for Hominidae (SYS85215) was found to be especially useful in restricting a search output to articles concerned with 'humans'. This BioSystematic code was always preferred to the use of the many synonyms for 'humans' (man, adolescent, girl, etc).

CROSS Codes: although there is no doubt as to the value of using these controlled language subject index terms in search profiles, several problems were encountered in their use:

(i) the CROSS code listing,(2) supplied by BIOSIS, is in code number order. It was necessary to prepare a CROSS code list in alphabetical order of subject as an aid to the liaison scientist in selecting the codes to be used in the search profile.

(ii) the CROSS codes correspond to relatively broad subject areas. Only 548 CROSS codes are available to cover the subject matter of the biological sciences. In contrast, Index Medicus uses approximately 7,000 separate subject headings for indexing the medical literature. In most cases, therefore, CROSS codes can only be used in coordinate searches involving the linkage either of two or more CROSS codes, or of CROSS codes with other term types (e.g. title words).

(iii) the detailed instructions provided to the BIOSIS indexers(3,4) are not available to BA Previews subscribers, and difficulties were encountered in recognising the subject area corresponding to some of the codes.

(iv) the subject headings of the CROSS code index are revised annually; some new headings are added and some existing headings are deleted or re-defined. All profiles must therefore be checked to discover whether amendments are needed as a result of changes in the CROSS index. Unfortunately, though the changes are implemented in the first BA Previews issue of each year, immediate revision of the profile is not possible as the "Guide to BA Previews"(6), which gives details of these changes, is not received by subscribers until March or April.

3. LITERATURE SEARCHING HABITS OF BIOLOGISTS

The most striking characteristic of literature usage by biologists, which was revealed by the initial questionnaire survey, was the great diversity of both the primary and the secondary sources which are used by the biological community. In all, 698 different primary journals were seen regularly by one or more of the 333 participants who completed the initial questionnaire. Only 66 of these journals were mentioned by ten or more users. The participants also cited 150 different secondary publications covering specialised subject areas (e.g. Water Pollution Abstracts), but only ten of these were mentioned by ten or more users.

Of the secondary publications which cover much broader subject areas, those most commonly used for current awareness searching were Current Contents (157 users, 47%), Biological Abstracts (105, 32%) and Chemical Abstracts (40, 12%). Only a minority of participants (22%) had previous experience of mechanised S.D.I. or retrospective search services and this experience most commonly took the form of a MEDLARS retrospective search.

The questionnaire replies provided detailed information regarding usage of the printed Biological Abstracts. It was found that only 178 (54%) of those who had volunteered to take part in the experiment had previously made any use of the printed publication for either current awareness or retrospective searches. Of the 154 who did not use it, 68 belonged to organisations which had no subscription. Of the 105 participants who used it for current awareness searches, only 33 claimed to use it regularly.

The most popular method of searching Biological Abstracts was item-by-item scanning of selected subject sections. This procedure was followed by 68% of those who used the journal for current awareness searches and by 64% of those who used it for retrospective searches. Though the majority also made some use of at least one of the four indexes (author, BASIC, BioSystematic, CROSS) there was evidence of substantial under-use of the printed indexes. In particular, only ten participants (6% of those who used Biological Abstracts) claimed to make use of the CROSS index. The BASIC index, which might have been expected to be extensively used for current awareness searching, was in fact used in only 49% of current awareness searches. In discussions with the participants, the impression was gained that a significant number of biologists either disliked the printed indexes or did not understand how they should be used.

4. SERVICE PERFORMANCE AND USER REACTION

The quantitative measures of performance, such as recall and precision, were based on user judgements as to the relevance of individual articles to their interests. The users' own opinions of specific aspects of service performance (e.g. precision, appearance of output), and also his overall judgement of the quality of the service were provided by his replies to the final questionnaire. The interpretation of these replies was often facilitated by informal discussions with participants. The measurement of performance parameters and the assessment of user reactions are therefore complementary aspects of an evaluation of a reference retrieval system.

The data which were collected in the present experiment defined the characteristics of an experimental BA Previews service. The significance of this data, as a means of predicting effectiveness and user satisfaction for a future permanent operational service, must be qualified by the fact that the experimental service has two characteristics which are unlikely to be present in a permanent BA Previews service:

(i) the use of liaison scientists to ensure intensive personal contact with users of the service. Their influence both on user interest and on quality of profiles has been discussed above (page 105).

(ii) the provision of a service in return for feedback but without any charge to the user. There were, therefore, none of the cost restrictions on the number of search terms and/or the size of output which are characteristic of the price structures of most cost-recovery or commercial services.

a) Output size, Recall and Precision

The mean output size of the 142 BA Previews profiles belonging to users who participated in the evaluation phase of the experiment and returned the necessary feedback was 34 items/issue. The estimated mean precision, i.e. the proportion of retrieved references which were judged relevant, was 38.6±1.9%. Approximately one third of the relevant items were judged by the user to be of MAJOR VALUE.

An estimate of 62.5±3.2% was obtained for the recall ratio - the proportion of relevant items in the data base which were actually retrieved by the search profile. As was found in the evaluations of other reference retrieval services(8,10,18), the mean recall for MAJOR value articles (66.4±3.8%) was significantly higher than that for MINOR value articles (55.9-4.4%).

Our precision and recall estimates are characteristic of the experimental service which was offered to users, and not of the data base alone. The values obtained are therefore influenced by a number of variables corresponding to the different parts of the retrieval system, i.e. the user-system interface, the search programs, the design of search strategies, and the data base. In particular, some recall and precision failures undoubtedly occurred as a result of inadequate communication between the user and the system. Thus, the search strategy may never even have attempted to retrieve some types of relevant material either because the user gave an incomplete description of his needs to the liaison scientist, or because the liaison scientist misinterpreted those needs. A detailed analysis of individual retrieval failures(8,9,10) would have been necessary in order to establish the relative proportions of precision or recall failures which are attributable to the communication process, and to each of the other subsystems which comprise the service.

b) Coverage

It was estimated that 75.4±2.1% of relevant articles (from serial publications)* encountered by the user in the course of his routine current awareness searching actually appeared in BA Previews, irrespective of whether or not the search profile was successful in retrieving these items. Samples of relevant references retrieved from three other major data bases were also used to estimate relative coverage ratios for three subsets of users whose interests were arbitrarily classified as biochemical, biomedical and "core biology" respectively. Each subset comprised approximately one third of the experimental population. For participants with some biochemical interests, the coverage ratio of *BA Previews*, relative to *CA-Condensates*, was 59.2±3.3%. For those with a biomedical bias, a coverage ratio of 72.3±2.8% relative to MEDLARS was obtained, and a sample of references retrieved from *Science Citation Index* gave a relative coverage ratio of 83.5±4.8% for the "core" biologists.

---

* All our coverage estimates are for articles from serial publications only. However, only the CAC sample contained a significant proportion (28%) of items from non-serial sources (books, government reports).

The majority of the articles which were not covered by *BA Previews* were from journals which were scanned by BIOSIS, but the individual articles had been excluded from the data base, presumably on the grounds of quality or because they were judged to be outside the subject area of the data base. With the exception of the *CAC* sample, over 80% of 'coverage failures' were of this type.

The *CAC* samples gave a lower relative coverage ratio than any of the other samples. This is almost entirely accounted for by the larger number of articles in the *CAC* sample which were from journals not scanned by BIOSIS. It appears that the journal coverage of *BA Previews* does not extend as far into biochemistry as it does into biomedicine.

We are not entitled to use these estimates to make *comparisons* between *BA Previews* and other data bases. Such comparisons would be possible only if estimates of the coverage of other data bases, relative to *BA Previews*, were also available. However, these results do suggest that a substantial proportion of biologists require information from more than one of the discipline-oriented services, and they therefore emphasise the need for research into the merging and repackaging of the existing magnetic tape data bases.

c) *Novelty Ratios*

An estimate of the novelty ratio of *BA Previews* showed that 76.9±1.8% of relevant items were not previously known to the user. The corresponding estimate for MAJOR VALUE items alone was 57.2±3.0%. It appears that, predictably, research workers are more likely to have prior knowledge of MAJOR VALUE than of MINOR VALUE articles. Evidence is presented in Chapter IV that the high value of the novelty ratio is a function of the retrieval performance of the *BA Previews* service, relative to the users' existing methods, rather than of the relative currency. The high value of the novelty ratio may partly be due to the fact that the users were less conscientious in their own manual searching as a result of receiving the *BA Previews* service.

However, a reduction in manual scanning would, in itself, be indicative of user confidence in the S.D.I. service.

d) Overall Recall

The novelty ratio of the service must be considered in conjunction with the estimated recall and coverage ratios. If the combined effect of recall and coverage performance is considered (the overall recall), it appears that the service was retrieving no more than 47% of the relevant material which was found by the user by his normal methods. On the other hand the novelty ratio (77%) indicates that the user was finding only 23% of the relevant material retrieved by BA Previews. Thus, to some extent, the BA Previews service complements rather than competes with the user's existing methods (though it need not be assumed that retrieval of all relevant items is necessary for complete user satisfaction).

e) Comparative Performance of BioResearch Index and Biological Abstracts

Outputs from issues of BioResearch Index (BRI) and of Biological Abstracts (BA) were almost identical with respect to the total number of items retrieved per issue (33.4 and 34.9 respectively) and the proportions both of MAJOR and of MINOR value articles in the output. The novelty ratio for relevant items in BRI (84.8±1.8%) was significantly higher than that for BA (73.4±2.1%). BioResearch Index contains items from non-serial publications such as US Government reports, books or conference proceedings, and also other material not considered suitable for inclusion in Biological Abstracts. The evidence presented above indicates that this type of material must not be considered as being of secondary importance to the user.

f) User Reactions to the Service

The two features of the service which appeared to be most appreciated were:

(i) it saved time by allowing users to reduce the time spent in searching the literature. This was mentioned as an important advantage of the service by 109 users (39% of the population). Furthermore, in the event of a permanent BA Previews service being available, 31% and 49% of users expected that they would be able to considerably reduce their scanning of current primary journals and of secondary publications respectively. Only 40 users (14%) did not think that a regular BA Previews service would enable them to make at least some reduction in their manual searching.

(ii) the service retrieved references which would not otherwise be located. This advantage was mentioned, in one form or another, by 92% of the population. This assessment by the users is supported by the high value obtained for the novelty ratio (77%) of relevant references.

The most frequently mentioned disadvantages of the service were:

(i) the delay between the publication of an article and its appearance in BA Previews (47% of the users).

(ii) the retrieval of too many irrelevant references (43%).

That the profile missed too many relevant references, or that the coverage was inadequate, were mentioned as major disadvantages of the service by only 9% and 10% of the users respectively.

A number of other disadvantages of the service, which did not relate to its retrieval performance or currency, were each mentioned by significant minorities: the difficulty of filing the computer printout (39%), no indication given of the language of the original paper (37%), the unattractive appearance of the computer printout (28%), and the absence of author addresses (27%) or of abstracts (27%) from the computer printout. Participants also expressed an overwhelming preference for card as opposed to paper output.

It appeared that a high level of user interest was maintained throughout the project. At the end of two years, 302* of the original 353 users were still receiving BA Previews printouts and returning feedback cards. In replies to the final questionnaire 194 users (69%) expressed an interest in subscribing to a permanent BA Previews current awareness service, though fewer (113) thought that the service was worth £50-£75. Replies to these two questions, which can be considered as expressions of the users' overall judgement of the quality of the service, were compared with replies to other questions relating to specific characteristics of the service. These comparisons revealed a pattern of factors which influenced the user's overall assessment, though no single characteristic exerted an overwhelming influence.

Users who nominated 'time-saving', 'convenient', or 'the retrieval of MAJOR VALUE references which would not otherwise be located' as important advantages, were significantly more likely to express an interest in subscribing to a permanent service and to judge that the service was worth £50-£75. Criticism of the retrieval of irrelevant material or, to a limited extent, of the lack of currency, tended to be associated with a negative reaction to the service. None of the other disadvantages mentioned by participants appeared to exert any influence on the user's overall judgement of the quality of the service. It can be concluded that these disadvantages are considered by the user to be annoyances rather than major drawbacks.

---

* This figure includes eight instances in which the original user had transferred his profile to a colleague.

Correlation of performance measures and questionnaire replies for individual users revealed that the novelty ratio was the best indicator of user satisfaction.

Participants appeared more satisfied with their BA Previews profile than they had previously been with the printed Biological Abstracts. This preference for the computer services was related, in part at least, to the under-usage of the printed indexes to which reference has already been made. In discussions with users, it frequently transpired that research workers had been unable to establish an effective method of searching the printed Biological Abstracts and were therefore failing to locate potentially interesting articles. As a result, some participants were pleasantly surprised at the number of relevant references retrieved by the profile. The latter offers a much less laborious means of searching the indexes and, in particular, the CROSS index which was rarely used for manual searches.

g) Interest in a Retrospective Search Service

Questionnaire replies revealed a considerable interest in a retrospective search service: 42% of participants considered that it would be of greater value than their S.D.I. service and a further 25% considered that it would have the same value.

h) The Information Centres

The eight information centres expressed a qualified interest in the service. Most centres appreciated that the service could provide supplementary material, usually amounting to approximately 10% of their total input. Their interest in subscribing to a permanent service depended on their assessment of the value of this additional material. They did not expect that the service would result in a significant saving of the time spent in manual searching as it would still be necessary for them to search the current journals in order to identify references as soon as possible.

5. FUTURE VIABILITY OF A BA PREVIEWS SERVICE

The experiment has shown that there is a substantial interest among biologists in a mechanised S.D.I. service and that BA Previews would be an acceptable data base for such a service. It was decided that the level of interest exhibited by participants justified the establishment, as a sequel to the present project, of a cost-recovery experiment at the University of Nottingham. This decision has been vindicated by the fact that 154 participants in the earlier project were willing to pay a £10 subscription, in addition to providing feedback, for their profile during the first year of the new experiment.* However, this charge is substantially below the full cost-recovery price of a BA Previews service which would probably be in the range £50-£75.

The extent of the demand for a permanent cost-recovery service and the viability of such a service will be determined by:

(i) the level of interest in the service among biologists

(ii) the willingness of organisations to make available the necessary funds

(iii) whether or not large industrial or government-funded research institutions elect to operate their own internal information services (possibly by purchasing the BA Previews tapes) rather than purchasing profiles from an outside organisation

The results obtained in the present project do not allow meaningful predictions as to the proportion of research biologists who would be willing and able to purchase a cost-recovery service during the next five years. In particular it must be noted that:

(i) the experimental population was biased in at least one respect: participants were likely to be enthusiastic consumers of information with an above-average interest in the development of new services

(ii) the reactions to the service, described in this report, are those of the individual research worker, who is unlikely to be responsible for the allocation of funds for these services. In future experiments, it may be desirable to assess the reactions of those who will have this responsibility.

---

*Enquiries about this new project should be addressed to Mrs J Stow, Department of Botany, University of Nottingham.

GLOSSARY OF TERMS USED IN THE EXPERIMENT

1. BioSystematic code: a 5-digit taxonomic code

2. CROSS code: a 5-digit subject code, which may be assigned at one of three levels. (See Primary, Secondary and Tertiary codes)

3. CRS: the prefix used to denote a CROSS code on the KDF-9 tape records

4. Enrichment term: keyword necessary to clarify or amplify the title - added at the end of the title by the BIOSIS editors

5. Evaluee: participant in the experiment who took part in the quantitative evaluation phase (where there is no possibility of misunderstanding, evaluees are simply referred to as 'users')

6. Feedback card: a card accompanying each computer printout on which the user was asked to mark the number of relevant references and write brief comments on the service, where applicable

7. Group profile: a profile which was written to satisfy the information needs of two or more workers, who attended the interview and contributed to the design of the profile.

8. IMPORTANT reference: "a reference directly related to your research interests, which you will read as soon as possible, i.e. it is essential that you know of it." (See MAJOR VALUE)

9. MAJ: the prefix used to denote a primary CROSS code on the KDF-9 tape records

10. MAJOR VALUE reference: "very centrally related to your research interests and could make a significant contribution."

11. MINOR VALUE reference: "of interest, and happy to have been informed of it, but of no great loss to have missed it."

12. NTX: the prefix used on the KDF-9 tape records to denote the BioSystematic code for an organism which belongs to a new or rearranged taxonomic group

13. Primary CROSS code: the CROSS code which corresponds to the printing assignment of the article, i.e. the subject heading under which it appears in the printed *Biological Abstracts*

14. RELEVANT reference: "a reference which, though not important, is of some interest to you, i.e. you are glad that you were notified of it."

15. S.D.I. (Selective Dissemination of Information): a personalised service, whereby individuals or groups are notified of the existence of current literature of potential interest. Their interests are described by means of a user profile.

16. Secondary CROSS code: a CROSS code used to indicate an important topic mentioned in the article, other than the printing assignment (Primary) code. (There may be one or more allocated to each article)

17. Shared profile: a profile which was written for one person, acting on behalf of one or more other workers

18. SUB: the prefix used to denote a secondary CROSS code on the KDF-9 tape records

19. SYS: the prefix used to denote a BioSystematic code on the KDF-9 tape records

20. Tertiary CROSS code: a CROSS code used to indicate a topic mentioned in the article which is of marginal interest. (There may be one or more allocated to each article)

21. User: participant in the experiment, who may or may not have taken part in the quantitative evaluation phase (includes all evaluees)

22. User profile (or Search profile): a collection of terms (i.e. words, codes, journal titles, authors' names) and of logical relationships among those terms, which define a user's interests in the form required for matching against a machine-readable data base

TIME-TABLE FOR THE EXPERIMENT

Date	Operation	Place	No. of Users
	1. **Development phase**		
1969			
1 December	M. Williams took up duties	Oxford	-
1970			
1 January	J. Stow took up duties	Nottingham	-
9 January	Seminar: recruitment for pilot study (M. Williams)	Oxford	16
26 January	B. Smith took up duties	Oxford	-
26-30 January	N.L.L. Course for Information Officers in University Libraries and others	Boston Spa	Project Staff
4 February	Seminar: recruitment for pilot study (J. Stow)	Nottingham	28
March	Completion of pilot studies	(Oxford (Nottingham	27 28
April	Visits to INSPEC, N.L.L. etc.		Project Staff
May 1970 - April 1972	Studies of data base:- 1. CROSS and BioSystematic codes 2. Enrichment terms	 Oxford Nottingham	 B. Smith J. Stow
May - August	Compilation of "Alphabetical index to CROSS Codes"		Project Staff
12 May	Advisory Committee	Oxford	
July-August	Visit to BIOSIS in the United States	U.S.A.	M. Williams
	2. **Operational phase**		
September-December	Main recruitment of users	-	353
26 November	Advisory Committee	Oxford	

TIME-TABLE FOR THE EXPERIMENT (CONT'D)

Date	Operation	Place	No. of Users
1971			
February-April	Second interviews	-	353
February	Reminder letter to users failing to return feedback cards	-	11 Deleted
April	Commencement of Evaluation		
7 May	Letter asking users if they would take part in Evaluation		342
18 May	Advisory Committee	Nottingham	
8-15 June	Distribution of BV-6 forms asking for references found by searching current literature (6-week period)	-	151 (Evaluees)
22 June	Last date for profile amendments	-	
28 June	Letter to users whose profile had been deleted, with questionnaire	-	59
July-August	Writing of Medlars profiles (BS) Writing of CAC profiles (JS) Design of Final Questionnaire (MW)	Oxford Nottingham Oxford	52 50
16 August	Final questionnaire (draft) Pilot study distribution	Oxford	55
1,9,15,28 September	Precision feedback: distribution of outputs representing the precision samples	Nottingham	151
23 September	Final questionnaire: distribution		302
October	Submission of SCI searches to ISI (MW)	Oxford	54
October-November	Final visits to users		270
12 October	Medlars printouts: distribution	Oxford	52
18 October	CAC printouts: distribution	Nottingham	50
15 November	Document samples: distribution	Oxford	151

TIME-TABLE FOR THE EXPERIMENT (CONT'D)

Date	Operation	Place	No. of Users
19 November	Advisory Committee	Oxford	
1 December	Output from SCI search: distribution	Oxford	54
1972			
January - May	Analysis and editing of feedback. Calculation of performance parameters		
March-June	Report writing	Oxford	
30 June	Submission of final draft of report to OSTI and members of advisory committee		
21 July	Advisory Committee	London	

BIBLIOGRAPHY

[Note: citations 1 - 7 are all published by BIOSIS (Philadelphia, USA). With the possible exception of reference 4 (no definite information available) all these publications are revised annually. The 1970-2 editions were used in the present experiment. All the BIOSIS publications are available to subscribers to BIOSIS services with the exception of references 3 and 4].

1. BIOSIS. A Guide to the Content of BA Previews.
[Contents include a description of the tape records, and details of the categories of enrichment terms used by BIOSIS, of the abbreviations used, and of additions, deletions or re-definitions of CROSS or BioSystematic codes].

2. BIOSIS. CROSS Code.
[A list in code number order of about 550 5-digit subject codes].

3.* BIOSIS. Section Definitions.
[An indexer's manual which gives detailed definitions of the subject areas covered by each CROSS code].

4.* BIOSIS. List of Priorities (1969).
[Instructions for the allocation of a primary CROSS code in cases where there are two or more alternatives, i.e. two or more subject areas are equally important to the main theme of the paper].

5. BIOSIS. BioSystematic Code.
[A listing of about 750 5-digit taxonomic codes, each of which represents a subdivision of the animal or plant kingdoms, according to phylum, class or group of related genera].

6. BIOSIS. Profiling Guide for *BA Previews*.
[Instructions for the construction and amendment of *BA Previews* profiles. Issued for the guidance of subscribers to the C.L.A.S.S. service offered by BIOSIS.]

7. BIOSIS. List of Serials.
[List of serial publications which are scanned by BIOSIS].

8. F.W. Lancaster. MEDLARS: Report on the Evaluation of its Operating Efficiency. *Amer. Doc.* *20*, 119-142 (April, 1969) and also: *U.S. Govt. Res. Rept.* PB 178 660.

---

* In-house publications not available to tape subscribers.

9. P. Leggate. The Evaluation of Operational Current Awareness Services: a Discussion of Practical and Theoretical Problems of Experimental Design. OSTI Report No. 5111. Experimental Information Unit, Oxford, 1971.

10. P. Leggate, B.N. Rossiter, J.F.B. Rowland. The Evaluation of an SDI Service Based on the *Index Chemicus Registry System*. *J. Chem. Doc.*, November, 1973 (in press).

11. UKCIS. Search Manual (1971).

12. Experiment with *BA Previews*. *OSTI Newsletter*, December 1969.

13. Literature Taped. *Nature*, *225*, 989-990 (1970).

14. F. Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bull.*, *1*, 80-83 (December, 1945); and also Probability Tables for Individual Comparisons by Ranking Methods. *Biometrics*, *3*, 119-122 (September, 1947).

15. S. Siegel. Nonparametric Statistics for the Behavioral Sciences. McGraw Hill Book Company Inc. 1956.

16. Westat Research, Inc. Procedural Guide for the Evaluation of Document Retrieval Systems. *U.S. Govt. Res. Rept.* PB 182 711 (December 1968), pp 137-185.

17. A. Kent, J. Belzer, M. Kurfeerst, E.D. Dym, D.L. Shirley, A. Bose. Relevance Predictability in Information Retrieval Systems. *Methods Inf. Medicine*, *6* (2), 46-51 (1967).

18. M.G. Corfield, R.J. Firth, G. Fraser, D. Hartley, P. Leggate, M.M. Norgett, J. Riley, B.N. Rossiter. The Liaison Scientist Experiment: A Study of the Provision of Mechanised Services to University Chemists. OSTI Report No. 5169, Experimental Information Unit, Oxford, May 1973.

APPENDIX II

CONTENTS

1. STUDY OF THE EDITORIAL PROCEDURES FOR ADDING KEYWORDS TO AUGMENT TITLES IN BA PREVIEWS

2. STUDY OF BIOSYSTEMATIC CODE USAGE

3. SPECIMEN PROFILE AND PRINTOUT FROM BA PREVIEWS

4. EXPLANATION OF BIOLOGICAL ABSTRACTS OUTPUT (1970)

APPENDIX II 1.

STUDY OF THE EDITORIAL PROCEDURES FOR ADDING KEYWORDS TO AUGMENT TITLES IN BA PREVIEWS

CONTENTS

A. INTRODUCTION

B. AIMS

C. IMPORTANCE OF ENRICHMENT TERMS IN SEARCHING BA PREVIEWS

D. METHODOLOGY

E. RESULTS
  1. Physical descriptors
  2. Organism(s): specific name, common name
  3. Geographical locations
  4. Drugs
  5. Drug affiliations
  6. Specific instrumentation

A. INTRODUCTION

As the BA Previews printout contains no abstract, the user must judge the relevance of the source document, in terms of his existing interests, from the title only.

In an attempt to make the title more meaningful, it is the accepted policy of BIOSIS to add keywords. No set thesaurus of terms has been compiled, so that this process involves the selection of natural-language terms by the editors. However, BIOSIS(1) gives some information concerning the categories of keywords which they include in their title enrichment process: "Keywords necessary to clarify or amplify the title are added at the end of the title in the following approximate order :-

- physical descriptor
- organism(s), specific name, common name
- organ system, or tissue used or affected

- geographical location, including geological strata
- drug(s) and affiliation(s) plus the word DRUG(S) linked to the name of a drug by a hyphen
- important chemicals
- specific chemical descriptor (affiliation)
- specific enzyme(s)
- specific instrumentation, apparatus, or methodology (when new, the word NEW is inserted ahead of the term it modifies)
- specific diseases
- purpose of experimentation"

B. AIMS

The purpose of this study was to examine some of these categories and determine the extent to which terms are added by the editors. From the point of view of profile construction it was considered important to determine the consistency with which certain categories of terms are added to ascertain if there was any control of synonyms. On the basis of the findings, it was hoped that some conclusions could be reached as to the degree of confidence which could be placed upon a significant term (e.g. hormone, drug, specific name of organism) not present in the title being added as an enrichment.

For example, a user who wishes to be alerted to references concerning the drug "Bethanidine" can only hope for 100% retrieval efficiency if he can rely upon "Bethanidine" always being added as an enrichment, if it is not present in the author title.

C. IMPORTANCE OF ENRICHMENT TERMS IN SEARCHING BA PREVIEWS

In a random sample of 100 references, it was found that an average of 3.1 words or phrases were added to the author title.

A sample of 150 references, which had been retrieved by word terms only, was selected from the output of five profiles. It was found that :

56% of the articles were retrieved by title words only
31% " " " " " " enrichment words only
13% " " " " " " either title or enrichment words

Enrichment terms are therefore responsible for a substantial proportion of the references retrieved.

D. METHODOLOGY

In this study, six categories of enrichment terms were studied:

Physical descriptors: a list of physical descriptors was compiled by scanning (a) the BASIC index of *Biological Abstracts* and enriched titles of *BioResearch Index* and (b) a computer-produced frequency list of enrichment terms used in six issues of *Biological Abstracts* and *BioResearch Index*.

Organisms, Geographical location and Drug names: these were investigated by scanning the abstracts in appropriate subject sections of the printed *Biological Abstracts* for the presence of these categories of terms; their addition as enrichment was then checked in the BASIC index. Checks were also made for the addition of geographical locations to titles in *BioResearch Index*.

Drug affiliations and Specific instrumentation: these were investigated using output from experimental profiles run against several issues of *BA Previews* (both *Biological Abstracts* and *BioResearch Index*).

E. RESULTS

1. Physical descriptors

The following terms which are used to describe publication types were found added as enrichment: review, abstract, letter, symposium, book, lecture, note, report, thesis, conference.

The descriptors "symposium" and "book" are added only when the complete publication is cited, and not to citations of individual papers or chapters (Fig. 1).

2. Organism(s): specific name, common name

BIOSIS(1) states that specific and common names are added "10 maximum in a series; if more are mentioned the individual taxa are grouped in larger categories". Table A shows samples of this type of enrichment.

It was found that the specific and common names of organisms are consistently added when they are not mentioned in the title but are referred to in the abstract (Tables B and C). If an article mentions the common name in the title and the specific name in the abstract, the latter would be included as an enrichment term (Table D). Conversely, if the specific name is mentioned in the title and the common name is in the abstract, it was found that the common name is not added as an enrichment term (Table E). If both the common name and specific name are referred to in the abstract, only the specific name, in general, is added as an enrichment term (Table F).

It is difficult to restrict a search to "humans" by using natural-language terms, as "human" could appear in the enriched title in numerous forms. The following have been observed: human, girl(s), boy(s), adolescent(s), adolescence, man, men, woman, women, infant, child, childhood, children, brother(s), patient, subject. For restricting a search to "human", the BioSystematic code for Hominidae (SYS86215) must be used.

Figure 1. ADDITION OF THE PHYSICAL DESCRIPTOR "SYMPOSIUM"

COSTANTINO I.N.
PROCEEDINGS OF THE LATIN AMERICAN CONFERENCE
ON THE CONSERVATION OF RENEWABLE NATURAL
RESOURCES SAN-CARLOS-DE-BARILOCHE ARGENTINA
MARCH 27 APRIL 2 1968 = SYMPOSIUM
INT UNION CONSERV NATURE NATUR RES ANNU REP
(13). 1968, 94-463
NUMBER 15981

(a) Complete publication: "Symposium" added as enrichment term

CRESPO J.A.
VAMPIRE AND OTHER BATS AND THEIR
CONSERVATION=DESMODUS-ROTUNDUS ECOLOGY
INT UNION CONSERV NATURE NATUR RES ANNU REP
(13). 1968, 102-104
NUMBER 15984

(b) Individual paper: "Symposium" not added as enrichment term

3. Geographical locations

BIOSIS(1) states: "A recognised geographical area (usually a country) is added at the end of the title for any geographical region occurring in the title or abstract".

a) Location given in the title

In general, if a specific city, town, county, river, mountain, estuary, lake, forest, research station, zoo, national park etc. is mentioned in a title, the corresponding country is added as an enrichment term.

Of 250 references which cited a geographical location of some type in the title, only 23 did not have the corresponding country inserted (Table G).

b) Location given in the abstract only

It was found that, when the geographical location is mentioned in the abstract but is not present in the title, it is not consistently added as an enrichment term.

Of 100 references which had a geographical location mentioned in the abstract, only 24 had the corresponding country added as an enrichment term (Table H). It appears that, if the site of the research is mentioned in the abstract, but is not considered to be an important feature of the work, the title is not enriched with the corresponding country. However, references were found where the location did not seem to be incidental to the research (Table I).

c) Conclusion

At present, it must be accepted that a profile cannot be effectively restricted to any particular geographical area, as it was not possible to find any standard procedure for this type of enrichment. Even if a location is mentioned in the abstract, it may not be added as an enrichment. Furthermore, there appears to be inconsistency regarding the choice of enrichment, e.g. references relating to a location in Great Britain could be enriched with Great Britain, Britain, British Isles, United Kingdom, England, Scotland or Wales. The form used tends to be that given by the author.

4. Drugs

BIOSIS[1] states that drugs are added "10 maximum in a series; if more are mentioned all drugs are grouped in larger categories". Table J shows examples which indicate that this statement is not always adhered to.

A sample of 77 references was examined of which 73 had all the drugs mentioned in the abstract inserted as enrichment terms. In a few cases drugs not mentioned in the abstract appeared as enrichment terms and must therefore have been mentioned in the original article.

There appears to be no consistent policy regarding the addition of drug compounds when trade names (or alternative names) are given in brackets. Table K shows that in some cases both are added; in some the former and in others the latter. BIOSIS[1] states that generic names are preferred to registered names.

5. Drug affiliation(s)

The policy regarding drug affiliations(1) is stated as follows:

1. For any drug name appearing in the title, the drug affiliation plus the hyphenated word DRUG are added once at the end of the title. Any drug name in the abstract is added at the end of the title and treated comparably. Generic names are preferred to registered names.

BEST COPY AVAILABLE

2. When a drug is discussed as having two different modes of action, each action is added.

3. The affiliations which are used relate to the specific action of the drug(s) as described or studied in the paper and not to the usual action associated with that drug.

4. The affiliative abbreviations listed below are based upon specific action categories for grouping drugs and conform to the drug classification of the United States Pharmacopeia, Seventeenth Revision (1965).

Anesthetic	ANESTHETIC-DRUG
antidote	ANTIDOTE-DRUG
antihistamine	ANTI HIST-DRUG
antiinfective	ANTI INFECT-DRUG
antiinflammatory agent	ANTI INFLAM-DRUG
antinauseant	ANTI NAUSEANT-DRUG
antineoplastic	ANTI NEOPLASTIC-DRUG
antiparasitic	ANTI PARASIT-DRUG
antiviral	ANTI VIRAL-DRUG
autonomic drug	AUTONOMIC-DRUG
carbonic anhydrase	CARBONIC ANHYDRASE INHIB-DRUG
cardiovascular	CARDIO VASC-DRUG
central depressant	CENT-DEPRESS-DRUG
central stimulant	CENT-STIM-DRUG
dermatological	DERMATOL-DRUG
diagnostic	DIAGNOS-DRUG
expectorant	EXPECTORANT-DRUG
gastrointestinal	GASTRO INTEST-DRUG
hematologic	HEMATOL-DRUG
hormone	HORMONE-DRUG
immunologic	IMMUNOL-DRUG
metabolic	METAB-DRUG
migraine (specific)	MIGRAINE-SPECIFIC-DRUG
oxytocic	OXYTOCIC-DRUG
pharmaceutical adjunct	PHARM-ADJUNCT-DRUG
radioprotectorant	RADIOPROTECT-DRUG
relaxant	RELAXANT-DRUG
renal acting	RENAL-ACTING-DRUG

The usage of the affiliations "HORMONE-DRUG" and "ANTI VIRAL DRUG" was investigated.

HORMONE-DRUG

A study was carried out to determine the reliability with which the drug affiliation "HORMONE-DRUG" is used to relate to the specific action of the drug(s). It was found that "HORMONE-DRUG" was consistently added if a particular hormone was used as a drug to observe its effect on a particular system.

ANTI VIRAL-DRUG

An experimental profile consisting of the CROSS code for anti-viral agents (CRS38506) and the term "ANTI VIR*" was run against five issues of *BA Previews* to determine the effectiveness of the drug affiliation "ANTI VIRAL-DRUG" compared with the CROSS code. The term "ANTI VIR*" was used in order to retrieve natural-language terms for anti-viral agents in the author title.

The results are shown in Table L. Whenever the drug affiliation "ANTI VIRAL-DRUG" was added as an enrichment term the CROSS code for anti-viral agents was also allocated. The references which were retrieved by "ANTI VIR*" and not by the CROSS code are shown in Table M. As articles can be retrieved by means of title words only, and not by the CROSS code or enrichment, it is not possible to rely exclusively upon the search term "ANTI VIRAL-DRUG" as a keyword in the profile.

It was also noted that an affiliation could be represented by several slightly different enrichment phrases, e.g.

ANTI VIRAL-DRUGS	52(05)No.24381
ANTI VIRAL DRUGS	52(06)No.32257
ANTI VIRAL-DRUG	52(06)No.32220
ANTI VIRAL	52(06)No.31389

Some references would also be missed if the affiliation alone were relied upon, as an agent effective against a virus might be considered to be primarily an IMMUNOL-DRUG [71(02)No.13843], an ANTI NEOPLASTIC-DRUG [71(02)No.13826] or an ANTI INFECT-DRUG [71(02)No.13806]. In such cases the enrichment ANTI VIRAL-DRUG would not be inserted and the references would only be retrieved by the CROSS code for anti-viral agents.

The CROSS code for anti-viral agents has a much broader coverage than

"ANTI VIR*" and retrieves general articles which are obviously relevant to the field, e.g.

> "UMEZAWA, H. Proceedings of the 6th International congress of chemotherapy. Section on chemotherapy of viral diseases [71(02)No.13822]."

It is apparent that, for complete coverage of anti viral agents, both the CROSS code and the natural-language term "ANTI VIR*" must be used.

6. Specific Instrumentation

In order to investigate the addition of instrumentation terms, a profile was constructed incorporating natural-language terms for "electron microscopy" and the CROSS code for "electron microscopy" (CRS01058).

The results shown in Table N indicate that the term "electron microscopy" is frequently added as enrichment when it occurs in the abstract. However, the grammatical form of the enrichment may differ from that used in the abstract.

Table P shows that, in order to retrieve references concerning "electron microscope studies", the enrichment term "electron microscope" alone is not sufficient. Enrichment terms used also include "electron micrograph" and "electron beam microscopy". The "electron microscopy" CROSS code (CRS 01058) especially at MAJ and SUB level, was usually assigned only to articles in which the technique itself was being studied and not to articles in which only the results obtained by the technique were of interest. For this type of profile all the synonyms for "electron microscopy" must be included in addition to the CROSS code for micro and ultra-microscopic anatomy (CRS 11108) when appropriate.

TABLES

A. Organisms: Examples of enrichment when more than 10 species are listed in the abstract.

B. Organisms: Examples of enrichment with common names of organisms mentioned in the abstract.

C. Organisms: Examples of enrichment with specific names mentioned in the abstract.

D. Organisms: Examples of enrichment when the common name is mentioned in the title and the specific name in the abstract.

E. Organisms: Examples of enrichment when both the common name and the specific name are mentioned in the abstract.

F. Organisms: Examples of enrichment when the specific name is mentioned in the title and the common name in the abstract.

G. Geographical location in the title: sample from 250 references.

H. Geographical location in the abstract: sample from 100 references.

I. Geographical location in the abstract: subject coverage of references.

J. Drugs: enrichment with drug names when more than 10 drugs are listed in the abstract.

K. Drugs: enrichment with drug names when trade names or alternative names are given.

L. Drug affiliations: comparison of retrieval by CROSS code (CRS38506) and natural-language term (ANTI VIR*).

M. Drug affiliations: retrieval by ANTI VIR* in the title and not by the affiliation ANTI-VIRAL DRUG or CROSS CODE (CRS38506).

N. Specific Instrumentation: relationship between terms mentioned in the abstract and terms added as enrichment.

P. Specific Instrumentation: comparison of retrieval by CROSS Code (CRS01058) and by natural-language terms.

TABLE A. ORGANISMS: EXAMPLES OF ENRICHMENT WHEN MORE THAN 10 SPECIES ARE LISTED IN THE ABSTRACT

BA VOL 51 ABSTRACT NO.	ABSTRACT	ENRICHMENT
28362	Boletus edulis, Boletus rufus, Boletus scaber, Lactarius rufus, L deliciosus, L torminosus, L rolemus, Birch mushrooms, orange mushrooms, aspen mushrooms, edible boletus, pepper mushrooms, vermin, pine forest, fir groves	Boletus spp. Lactarius spp. Pine, fir, vermin
23933	quotes 7 individual species plus 5 Polygonum species	7 species Polygonum-spp
34714	approximately 40 species listed	-
34695	quotes 14 Culicoides	14 new record species
40473	species of Unionidae and Margaritiferidae listed	Amblemella new genus Denminaia new genus Hemisolasma Parreysia Lamellidens Limnoscapha Nannonaia Trapezoideus Rectidens, Physunio Unio Pseudodon Monodonlina Leguminaia Cosmopseudon 28 new species 2 new forms fossil

TABLE B. ORGANISMS: EXAMPLES OF ENRICHMENT WITH COMMON NAMES OF ORGANISMS MENTIONED IN THE ABSTRACT (SAMPLE FROM 125 REFERENCES)

BA VOL 51 ABSTRACT NO.	TITLE	ABSTRACT	ENRICHMENT
45809		wheat, barley soybean, sweet potato groundnut, rice	wheat, barley soybean, sweet potato groundnut, rice
35110	-	man, ants	man, ants
35138	mammalian	mouse, fibroblasts spinach, African violets, chicken liver	mouse, fibroblasts spinach, African violet, chicken
35137		human embryonal lung strain (HELS)	human embryonal lung strain
35145	birds	chickens, guinea fowl	chickens, guinea fowl
35446	ferrets	rats, cats, possums	rat, cat, possum
35478	-	male and female	human
35479	neonate	infant	human
37947	-	rat liver	rat
37965	-	rats	rats
37971	-	mice	mouse
37972	-	person, patient	human
37977	-	rats	rat
37997	-	mice	mouse
38001	-	human	human
38017	-	carrot	carrot
38021	fruits	apples, plums, peas, cherries, wheat	apple, plum, pea, cherry, wheat
38024	legumes	soybeans	soybeans

TABLE C. ORGANISMS: EXAMPLES OF ENRICHMENT WITH SPECIFIC NAMES MENTIONED IN THE ABSTRACT (SAMPLE FROM 42 REFERENCES)

BA VOL 51 ABSTRACT NO.	ABSTRACT	ENRICHMENT
35110	Hippuris vulgaris	Hippuris-vulgaris
35169	C. blumei, C. barabatus	Coleus-blumei Coleus-barabatus
35180	Glycine max	Glycine-max
29729	Plantago lanceolata Rumex acetosa Dactylis glomerata	Plantago-lanceolata Rumex-acetosa Dactylis-glomerata
29740	Cryptostegia grandiflora Pergularia daemia Calotropis procera Leptadenia pyrotechnica	Cryptostegia-grandiflora Pergularia-daemia Calotropis-procera Leptadenia-pyrotechnica
29750	Thalaspi alpestre Minuartia verna	Thalaspi-alpestre Minuartia-verna
29756	Deroceras reticulatum Arion fasciatus A. subfuscus A. intermedius A. hortensis D. caruanae	Deroceras-reticulatum Arion-fasciatus Arion-subfuscus Arion-intermedius Arion-hortensis Deroceras-caruanae
34550	Camerina-catenuia Eoconuloides parvulus Eoconuloides wellsi Helicostegina polygyralis Helicolepidina spiralis	Camerina-catenuia Eoconuloides-parvulus Eoconuloides-wellsi Helicostegina-polygyralis Helicolepidina-spiralis
33622	Biomphalaria tenagophilia Schistosoma mansoni	Biomphalaria-tenagophila Schistosoma-mansoni
23923	Pistacia lentiscus Quercus ilex	Pistacia-lentiscus Quercus-ilex

TABLE D. ORGANISMS: EXAMPLES OF ENRICHMENT WHEN THE COMMON NAME IS MENTIONED IN THE TITLE AND THE SPECIFIC NAME IN THE ABSTRACT

BA VOL 51 ABSTRACT NO.	TITLE	ABSTRACT	ENRICHMENT
24020	Striped skunk	Mephitis mephitis	Mephitis-mephitis
24021	Racoon	Pyocyon lotor	Pyocyon-lotor
41320	Snapping turtle	Chelydra serpentina	Chelydra-serpentina
41322	White bass	Roccus chrysops	Roccus-chrysops
35446	Ferrets	Putorius putorius furo	Putorius-putorius-furo
35454	Black headed gull	Larus ridibundus	Larus-ridibundus
39757	Bluegull	Lepomis macrochirus	Lepomis-macrochirus
38436	Human	Homo sapiens	Homo-sapiens
29723	Leafy liverwort	Scapania nemerosa	Scapania-nemerosa
28348	Pear	Pyrus ussuriensis	Pyrus-ussuriensis
28398	Taiwan red and yellow cyprus	Chamaecyparis formosensis Chamaecyparis taiwanesis Chamaecyparis obtusa	Chamaecyparis-formosensis Chamaecyparis-taiwanesis Chamaecyparis-obtusa
23292	Great crested grebe	Podiceps cristatus	Podiceps-cristatus
23319	Dolphin	Delphinus delphis	Delphinus-delphis
17789	Harbor seal	Phoca vitulina	Phoca-vitulina
18032	Tawny owl	Strix aluco	Strix-aluco
12193	Dogwood	Cornus florida	Cornus-florida
16442	Rough cinquefoil	Potentilla norvegia	Potentilla-norvegia
16699	Clover	Trifolium subterraneum	Trifolium-subterraneum
16919	Tobacco	-	Nicotiana-tabaccum
16948	Pineapple disease	Ceratocystis paradoxa	Ceratocystis-paradoxa

TABLE E. ORGANISMS: EXAMPLES OF ENRICHMENT WHEN BOTH THE COMMON NAME AND THE SPECIFIC NAME ARE MENTIONED IN THE ABSTRACT

BA VOL 51 ABSTRACT NO.	ABSTRACT	ENRICHMENT
24026	Field vole, Microtus arvalis	Microtus-arvalis
23972	Fleas, Amphalius-runatus Ctenophyllus bondari	Amphalius-runatus Ctenophyllus-bondari
47156	Great horned owl Bubo virginianus,Long eared owl, Asio-otus, Burrowing owl, Speotyto cunicularia Barn owl, Tyto alba	Bubo-virginianus Asio-otus Speotyto-cunicularia Tyto-alba
62811	Lettuce, Lactuca sativa	Lactuca-sativa
35022	Sweet chestnut Castanea sativa	Castanea-sativa
41273	Fagus, Magnolia, Quercus, Carya, Liquidambar, Pinus, Illicium floridanum Beech, magnolia, oak, hickory, sweetgum, pine	Fagus, Magnolia, Quercus, Carya, Liquidambar, Pinus Illicium-floridanum
41327	Yellowfin tuna Thunnus albacares Bigeye tuna T. obesus	Thunnus-albacares Thunnus-obesus
38024	Kidney beans Phaseolus vulgaris	Phaseolus-vulgaris
29789	Salmo trutta Brown trout, Brook trout, Salvelinus fontinalis	Salmo-trutta Salvelinus-fontinalis
29802	Bay mussel, Mytilus edulis fluffy sculpin, Oligotocottus synderi	Mytilus-edulis Oligotocottus-synderi
34355	Chestnut blight Endothia parasitica	Endothia-parasitica
34876	Blowfly, Calliphora phaenicia Housefly, Musca domestica	Calliphora-phaenicia Musca-domestica
23298	Olive fly, Dacus oleae	Olive fly, Dacus-oleae
18070	Coho salmon, Oncorphynchus kisutch	Oncorphynchus-kisutch
12150	Lololly Pine, Pinus-taeda	Pinus-taeda
17390	Crayfish, Panulirus longipes	Panulirus-longipes
58224	Skipjack tuna, yellowfin tuna, bluefin tuna, Katsuwonus pelamis Thunnus albacares Thunnus thynnus	Katsuwonus-pelamis Thunnus-albacares Thunnus-thynnus

TABLE F. ORGANISMS: EXAMPLES OF ENRICHMENT WHEN THE SPECIFIC NAME IS MENTIONED IN THE TITLE AND THE COMMON NAME IN THE ABSTRACT

BA VOL 51 ABSTRACT NO.	TITLE	ABSTRACT	ENRICHMENT
23929	Heterpogon contortus	black spear grass	-
	Themeda australis	kangaroo grass	-
41281	Quercus turbinella	live oak	-
	Quercus emoryi	emory oak	-
41305	Shaerium corneum	fingernail clam	-
	Bufo vulgaris	toad	-
29720	Ceanothus velutinus	snowbush	-
34871	Tribolium casteneum	redrust flour beetle	-
	Tribolium confusum	confused flour beetle	-
19539	Tritusus pyrrhogaster	Japanese newt	-
22906	Lissocarcinus echinodisci	crab	-
40994	Artemia salina	brine shrimp	-
34815	Polinices lewisi	clam drill	-
34990	Eutamias quadrivittatus	chipmunks	-
34035	Hordeum vulgare	barley	-
34078	Glycine max	soy beans	-
34091	Lemna gibba	duckweed	-
39781	Theobroma cacao	cacao tree	-
39812	Allium cepa	onion	-
39962	Pinus edulis	pinyon pine	-
	Pinus albicaulis	whitebank pine	-
22229	Pisum sativum	peas	-
22422	Cynodon dactylon	Bermuda grass	-
22429	Elymus junceus	wild rye	-
22684	Plutella maculipennis	diamond black moth	-

TABLE G. GEOGRAPHICAL LOCATION IN THE TITLE : SAMPLE FROM 250 REFERENCES

ISSUE	ABSTRACT NO.	LOCATION IN TITLE	LOCATION AS ENRICHMENT
70/01	1010	Newfoundland	Canada
	5123	Arkansas	USA
	3526	Newfoundland	Canada
	2885	Colorado Mountain	USA
	1837	Lake Champlain	-
	1017	Norbotten county	-
	2818	Colorado	USA
	6125	River Axe Devon	England
	5098	Mackenzie and Coppermine Rivers	Canada
	5032	Niagara	New-York USA
	6061	Spietbergen	Norway
51/01	5825	California	USA
	520	N.W. Sutherland	England
	550	Volga River	USSR
51/02	6507	Azov Sea	USSR
	6131	Lake Maggiore	Italy
	6317	Amazon	Rio-Negro
	5866	Amazon	Brazil
	6110	Hooghley Estuarine	India
	6133	Lake Suwa	Japan
	6155	Lake Hibara	Japan
	6310	N. Moravian Region	-
	11162	British Columbia	Canada
	10201	Alberta	Canada
	11072	Wisconsin	USA
	6530	Alberta	Canada
	6377	N.W. Moravia	Czechoslovakia

TABLE H. GEOGRAPHICAL LOCATION IN THE ABSTRACT:
SAMPLE FROM 100 REFERENCES

BA VOL 51 ABSTRACT NO.	ABSTRACT	ENRICHMENT
23952	Derbyshire	-
35520	South California	California USA
18014	Southern Sweden	-
18147	Africa, Uganda, Kenya South Africa	-
18160	Pisa Florence	-
35508	Colorado desert	-
35510	Wisconsin	-
35583	Marion Lake	-
18021	Africa, Australia, America	-
40230	England, Scotland, Wales, Carpathian Provenances	United Kingdom Czechoslovakia
28605	Mergab River Valley	USSR
727	Holland, USSR, N. Africa, S. Europe	-
35536	Nova Scotia, New Brunswick	-
58751	Oregon coast	-
64655	Georgia	-
64657	Kentucky	-
64665	Oklahoma	-
64668	Nevada	-
64670	Michigan	-
64671	Newport River Estuary S. Caroline	-
64678	Texas, Mexico	-
47137	W. Hungary, E. Poland, N & S German Hill Country	-
47162	Toronto	-
47169	Swedish Lake	-
47174	Nanao Bay Japan	Japan

TABLE I. GEOGRAPHICAL LOCATION IN THE ABSTRACT:
SUBJECT COVERAGE OF REFERENCES

BA VOL 51 ABSTRACT NO.	ABSTRACT	ENRICHMENT	SUBJECT COVERAGE
64657	Kentucky	-	Ecological study of a plant species to explain its local distribution
64668	15 sites in Nevada	-	Study of desert rodents
47137	W. Hungary E. Poland N & S German Hill country	-	Forest associations from a geographical viewpoint
58751	Oregon Coast	-	Water table depths and vegetation depths measured at Oregon
58783	S. Ontario	-	Migration of starlings from S. Ontario
58721	Canada	-	Analysis of wind speeds at forestry stations and airports across Canada
47128	Isle of Man County Cork	-	Comparison of plants in different habitats

TABLE J. DRUGS: ENRICHMENT WITH DRUG NAMES WHEN MORE THAN 10 DRUGS ARE LISTED IN THE ABSTRACT

BA VOL 51 ABSTRACT NO.	ABSTRACT	NO ENRICHMENT TERM*
136538	chlorpromazine, trifluoperazine, perphenazine, thioridazine, haloperidol,ethomoxane, chlorprothixene, tetrabenazine, reserpine, benztropine, chlordiazepoxide, pentobarbital, meprobamate, chloral hydrate, paraldehyde, scopolamine, cyproheptradine, atropine, trihexyphenidyl	
37856	phenelzine, iproniazid, iso carboxazide, imipramine, harmaline, pheniprazine, pargyline, tranylcypramine, amitriptyline, cocaine, amphetamine, DL-DOPA, 5-hydroxtryptophan	imipramine amitriptyline
49258	adumbran, librium mogadon, valium, tegretal, tofranil, pertofran, insidon, noveril, laroxyl, acetexa	
49370	papaverine, euphyllin, nitroglycerine, chloracizine, chlorpromazine, reserpine, hexamethanium, morphine, promedol, fibrinolysin, isonicotinic acid, ipiazide	

---

* In all cases except for the two compounds named, the drug name appears as an enrichment term.

TABLE K. DRUGS: ENRICHMENT WITH DRUG NAMES WHEN TRADE NAMES OR ALTERNATIVE NAMES ARE GIVEN

BA VOL 51 ABSTRACT NO.	TITLE	ABSTRACT	ENRICHMENT
136513		Eserine (Physostigmine)	Eserine Physostigmine
136502		NSD1015X (m-hydroxy benzyl hydrazine)	m-hydroxy benzyl hydrazine
138121		Arzenobenzene (Arzene),	Arzene benzine
		Crystal violet,	Crystal violet
		0,0 dimethyl-2-2-2 tri chlorohydroxymethyl-phosphoro acid ester (Neguvon),	Neguvon
		Halquinol (Quixalud),	Halquinol
		Glycolnaisul phosphate (Plewin),	Glycolnaisul phosphate
		Sulfaquinoxaline sodium (Noxal),	Sulfaquinoxaline sodium
		Quinacrine hydrochloride (Atobrine),	Quinacrine hydro chloride
		Neomycin sulfate and Oxytetracycline (Neo-terra),	Neomycin sulfate Oxytetracycline
		Formalin, Whitsyn-S	Formalin Whitsyns
37711		Alpha-acetydigoxin (Sandolania)	Alpha acetydigoxin
37886	Dibenzepin	(Noveril)	Noveril
	Impramine	(Tofranil)	Tofranil
87905	Oxytocin	(Syntocinon)	Syntocinon
37926	Trilafon	(Perphenazine)	Perphenazine

TABLE L. DRUG AFFILIATIONS: COMPARISON OF RETRIEVAL BY CROSS CODE (CRS38506) AND NATURAL-LANGUAGE TERM (ANTI VIR*)

ISSUE	RETRIEVED BY BOTH ANTI VIR* AND CRS38506	RETRIEVED BY ANTI VIR* ONLY	RETRIEVED BY CRS38506 ONLY
71(02)	40	3	17
52(03)	4	-	3
52(04)	4	1	1
52(05)	13	-	6
52(06)	11	2	6

TABLE M. DRUG AFFILIATIONS: RETRIEVAL BY ANTI VIR* IN THE TITLE AND NOT BY THE AFFILIATION ANTI VIRAL-DRUG OR CROSS CODE (CRS38506)

Abstract No.	Terms in Title
52(04)20139	anti viral antibody
52(06)31635	anti viral antibody
52(06)32246	anti viral activity
71(02)10700	anti viral action
71(02)100314	anti viral resistance
71(02)13983	anti viral substance

TABLE N. SPECIFIC INSTRUMENTATION: RELATIONSHIP BETWEEN TERMS MENTIONED IN THE ABSTRACT AND TERMS ADDED AS ENRICHMENT

BA VOL 52 ABSTRACT NO.	CRS01058	TITLE	ENRICHMENT	ABSTRACT
128151	✓	Fine structure	-	Electron microscope
128180	✓	" "	-	Electron microscope
128217	✓	" "	-	Electron microscope
127090	✓	" "	-	-
126130	✓	" "	-	-
126127	✓	" "	-	-
128173	✓	" "	Electron microscopy	Electron microscope
128190	✓	" "	Electron microscopy	Electron microscope
128252	✓	" "	Electron microscope	Electron microscope
127126	✓	" "	Electron microscopy	Electron microscopy
121071	✓	" "	Electron microscope	Electron microscopic
123703	✓	" "	Electron microscope	Electron microscope

TABLE P. SPECIFIC INSTRUMENTATION: COMPARISON OF RETRIEVAL BY CROSS CODE (CRS01058) AND BY NATURAL-LANGUAGE TERMS

ISSUE	CRS01058 ONLY	CRS01058 + ELECTRON MICRO*	CRS01058 + ULTRA STRUCT*	CRS01058 + FINE STRUCT*	ELECTRON MICRO* ONLY	ULTRA STRUCT* ONLY	FINE STRUCT* ONLY
71(10)	2	97	1	-	3	8	52
71(02)	10	90	15	3	1	5	22
71(03)	2	115	21	5	1	7	25
52(03)	22	123	23	2	5	18	20

APPENDIX II 2.

STUDY OF BIOSYSTEMATIC CODE USAGE

CONTENTS

A. METHODOLOGY

B. RESULTS

1. Coding omissions
2. Coding errors
3. Studies of BioSystematic code allocation in six areas of applied biology
4. Scope of very general BioSystematic codes
5. Overlap between CROSS and BioSystematic codes

C. RECOMMENDATIONS FOR USING BIOSYSTEMATIC CODES

1. Phylogenetic considerations
2. Subject range of BioSystematic codes

A. METHODOLOGY

Studies of the allocation of BioSystematic codes were carried out, using specially constructed profiles which were run against one or more issues of BA Previews. The output from these experimental profiles included a printout of the codes allocated by BIOSIS to each paper. For some aspects of the study, it was also necessary to scan the corresponding abstracts in the hard copy Biological Abstracts but, with one exception, reference was not made to the original article.

B. RESULTS

1. Coding Omissions

In order to determine whether BioSystematic codes are always assigned when an organism is mentioned in either the title or the abstract, word terms and taxonomic CROSS codes were used to construct profiles corresponding to seven taxonomic groups. Table A shows the results of an analysis of the articles retrieved by each of the seven profiles.

TABLE A. CODING OMMISSIONS

Topic and appropriate SYS codes	Word terms and CROSS codes used in profiles	BA Previews Issues	No. of refs retrieved	No. of refs with correct SYS code	No. of refs without SYS code
1. ECHINODERMS SYS83	Holothuroid* sea slug* Ophiuroid* Asteroid* Echinoderm* Echinoid* brittle star* Crinoid* star fish* sea urchin* heart urchin* sea lil* cake urchin* sea star* sea cucumber* CRS64048 feather star* CRS63548	51(09) 51(10) 52(01) 52(02)	57	57	-
2. VIRUSES SYS03	virus* viral *phage*	51(09)	187	183	4
3. ALGAE SYS13 SYS14	CRS50504(General/Systematic)	51(09)	14	14	-
4. FUNGI SYS15	CRS50506(General/Systematic)	51(09)	54	54	-
5. CONIFERS SYS251	pine* Pinus conifer* CRS50518(General/Systematic)	51(09)	50	50	-
6. AMPHIBIA & REPTILIA SYS853 SYS854	CRS62512(Amphibia/Reptilia) CRS62514(Amphibia) CRS62516(Reptilia)	52(01) 52(02) 70(12)	53	53	-
7. PIGS SYS85740	pig* hog* porcine Sus sow* pork boar* swine	51(09)	96	96	-

NB "Ignore" terms have been omitted from the table for nos. 2, 5 and 7

Asterisks (*) denote truncation signs

The four articles in which an organism was mentioned in the enriched title, but the appropriate BioSystematic code was not allocated are listed below.* In each case, the organism for which a code was not assigned is indicated by underlining.

"Electron microscopic observations on the association of viruses with membrane systems in hamster tumor cells propagated in tissue culture" (51, No.9, Abs.49992)

"Attempt to increase the effect of regional chemotherapy in experimental tumors by procaine" = virus transformed rate fibroblasts exposed to thymine
(51, No.9, Abs.50008)

"Antibodies to Epstein Barr virus in Burkitt's lymphoma and control groups" (51, No.9, Abs.49840)

"Malignant and transforming activity of Rous sarcoma virus. Part 1. Malignant effects of Rous sarcoma virus"
(51, No.9, Abs.49986)

However, the appropriate BioSystematic code was allocated in all the other 447 articles which were examined. It can therefore be concluded that, if an organism is mentioned in the enriched title and/or the appropriate taxonomic CROSS code is allocated, the appropriate BioSystematic code for the organism is assigned almost invariably.

2. Coding Errors

One or more BA Previews tapes were searched for all articles containing the BioSystematic codes for five taxonomic groupings. The search profiles made use of the partially hierarchical structure of the BioSystematic codes. Thus the search term SYS03 would retrieve articles which had been assigned any of the following codes:

SYS03000 (Viruses - general); SYS03200 (Animal viruses);
SYS03400 (Plant viruses); and SYS03600 (Bacterial viruses)

The title and abstract appearing in the printed Biological Abstracts were consulted for each of the retrieved articles in order to discover whether the BioSystematic code had been assigned appropriately. The results are shown in Table B.

---

* Words following the equality symbol (=) are enrichment terms added by the BIOSIS indexers

In all but one of the articles examined the BioSystematic code corresponded to an organism mentioned in either the title or the abstract. The single exception was:

> "Vineyard sprayers' lung. A new occupational disease" = guinea pigs copper sulfate hydrate lime toxicity
> (51, No.9, Abs.49735)

The article was assigned the BioSystematic code for Echinodermata in both the magnetic tape and printed indexes, but neither the abstract nor the original paper contained any reference to this type of organism. This appeared to be an isolated clerical or indexer error. The probability of a BioSystematic code for a particular taxonomic group retrieving an article unrelated to that group is therefore negligible.

Table B also shows that the BioSystematic code corresponded to an organism mentioned in the title in 91% of the articles examined. The BioSystematic index therefore appears to be only slightly more exhaustive than the enriched title. However, the codes would usually be preferred to natural-language terms, as with the latter allowance must be made for synonyms and alternative grammatical forms.

3. Studies of BioSystematic code allocation in six areas of applied biology

It would be expected that in articles concerned with applied biology any organisms mentioned would often be of secondary importance to the main theme of the paper. The consistent allocation of BioSystematic codes to such articles is obviously an advantage to some applied biologists, but it was also found to be a disadvantage for users who were interested in specific organisms but not in relation to applied biology. For some users in this category, the use of BioSystematic codes retrieved an excessive number of irrelevant items.

A study was therefore made of the allocation of BioSystematic codes to articles from six fields of applied biology. A sample of articles from each field was retrieved by search profiles containing either CROSS codes or word terms as shown in Table C.

TABLE B. CODING ERRORS

Topic	SYS code	BA Previews Issue	Total no. of refs retrieved	No. of refs where title appropriate	No. of refs where abstract appropriate	No. of refs where neither appropriate
1. VIRUSES	SYS03	51(09)	263	232	31	-
2. ALGAE	SYS13/14	51(09)	87	81	6	-
3. CONIFERS	SYS251	51(09)	85	75	10	-
4. PIGS	SYS85740	51(09)	99	96	3	-
5. ECHINODERMS	SYS83	51(09) 51(10) 52(01) 52(02)	65	62	2	1
TOTALS			599	546	52	1

TABLE C. ALLOCATION OF BIOSYSTEMATIC CODES IN APPLIED BIOLOGY

Topic	Terms & codes in profile	BA Previews Issue	Total no. of refs retrieved	No. of refs with SYS codes allocated	No. of refs with SYS codes not allocated
1. FOOD TECHNOLOGY	CRS13504(Fruit & Vegetables) CRS13510(Cereal Chemistry) CRS13522(Fish & Freshwater Products)	70(06) 51(12) 51(14)	111	103	8
2. FORESTRY AND FOREST PRODUCTS	CRS53500	51(09)	158	150	8
3. PLANT-DERIVED DRUGS	ALKALOID* MORPHINE DIGITALI* COLCEMID* COLCHICINE EPHIDRINE	52(02)	37	35	2
4. MYCOLOGY AND ANTIFUNGAL AGENTS	CRS36008(Mycology) CRS38508(Antifungal Agents)	51(09)	40	35	5
5. TISSUE CULTURE	CRS32500	51(09)	139	137	2
6. CLINICAL MEDICINE	CRS00531(Forensic Science) CRS19006(Dental & Oral Pathology) CRS12502(General Pathology) CRS12503(Comparative Pathology) CRS21004(Addiction)	51(09)	112	110	2
TOTALS			597	570	27

a) Food Technology

References concerning fruit, vegetables, cereals and fish were retrieved by these three CROSS codes in the context of:

(i) processing as items of food

(ii) nutritive qualities

(iii) chemical composition of food

(iv) spoilage and contamination.

Of 111 articles retrieved, only 8 had no BioSystematic code corresponding to the organism acting as the source of food and five of these did not mention the type of food involved. The remaining three articles to which the BioSystematic code was not assigned were:

"Vitamin B-6 components in some meats, fish, dairy products and commercial infant formulas"
(51, No.12, Abs.65300)

"Overweight and dietary habits of adolescents = vegetable fats butter cheese meat fish"
(51, No.14, Abs.76751)

"Lycine enrichment of wheat in flour evaluation in infants"
(51, No.14, Abs.76733)

b) Forestry & Forestry Products

A profile consisting of the CROSS code 53500 retrieved articles relating to trees and other plants in the context of:

(i) forest fire risk

(ii) regional distribution of various species

(iii) timber products e.g. wood, sawdust, resin

(iv) climatic conditions

(v) afforestation programs

(vi) soil requirements and nutrition

(vii) forest pests and forest animals.

Of the 158 references retrieved, only 8 had no appropriate BioSystematic code and none of these mentioned specific genera or species.

c) Plant-derived Drugs

A title-word search for drugs of plant origin retrieved papers concerned with:

(i) therapeutic medicine

(ii) tissue culture

(iii) chemical analysis

(iv) chemical extraction

(v) pharmacological experiments on animals.

All but two of the 37 retrieved articles were indexed with the appropriate BioSystematic code for the plant from which the drug derived.

d) Mycology and Antifungal Agents

References to fungal species were retrieved by the CROSS codes in the context of:

(i) disease-carring pathogens

(ii) diagnosis of fungal infection

(iii) drug treatment of fungal diseases.

All but five of the 40 references retrieved had a fungal BioSystematic code allocated. Of these five only two made even a slight reference to "mycotic infection" or "antifungal treatment"; the remaining three made no reference at all.

e) Tissue Culture

This CROSS code retrieved articles concerned with:

(i) irradiation of animal cells

(ii) cell culture media and methods

(iii) cells used to culture viruses and bacteria

(iv) pharmacological response of cells in vitro

(v) metabolism of cells in culture.

In all but two cases, the article was indexed with the BioSystematic code for the animal or plant from which the cells were originally obtained.

The two exceptions were;

"The in vitro transformation by an avian adenovirus" = *chick* embryo lethal orphan virus
(*51*, No.9, Abs.49995)

"Correlation between immunoreactive growth hormone and prolactin activity in human and simian pituitary cell cultures" = fetus *pigeon* crop sac method
(*51*, No.9, Abs.50212)

f) Clinical Medicine

Articles on human clinical medicine retrieved by these CROSS codes included such terms as:

human(s)
man/men
woman/women
mother(s)
female(s)
child(ren)
infant(s/ile)
adult(s)
adolescent(s)
aged, veteran(s)
Indian(s)
hipp(y/ies)
patient(s)
case(s)
clinical
addict(s)

alcohol(ism/ic(s))
schizophren(ia/ic(s))
medico-legal
forensic
paternity
legislation
public health
rehabilitation
drunkenness
teenage(rs)
girl(s)
physically disabled
medical education
bab(y/ies)

and many other terms, all indicative of some aspect of human medicine. Of the 112 references retrieved, only two had not been given the BioSystematic code for *Homo sapiens*. These were:

"Scope of immunology in science, research and *clinical* terms"
(*51*, No.9, Abs.50644)

"The biologic effects of UV radiation with emphasis on the skin" = book symposium *erythema* *phytoallergy* phyto-toxicity
(*51*, No.9, Abs.46690)

4. Scope of very general BioSystematic codes

Several very broad categories are included in the BioSystematic Code Index, e.g.

00500	Organisms (General)
01000	Micro-organisms (General)
11000	Plantae (General)
33000	Animalia (General)
34000	Invertebrata (General)
85150	Vertebrata (General)
85700	Mammalia (General)

The BioSystematic index is not fully hierarchical and a search for one of these general codes will not retrieve articles which have been given only one of the more specific codes corresponding to the subdivisions of these phylogenetic categories. It is therefore necessary to know when and how these general codes are assigned. If BIOSIS(1) adheres to the stated editorial policy of assigning the most specific of the available codes, the general codes should only be used for articles which describe general principles and/or fail to mention specific organisms. 243 references retrieved by the codes for Micro-organisms, Chordata and Vertebrata were classified as shown in Table D, from which it will be seen that:

(i) 186 references did not mention any specific organisms though some of these referred to 'microbes', 'fauna', etc. which were not otherwise specified.

(ii) 29 references dealt with general principles though a few animals were mentioned specifically.

(iii) 28 references did not appear to deal with general topics and referred to only one or a few species. The allocation of a more general BioSystematic code would appear to be a coding error.

Since the two categories, Chordata and Vertebrata, could overlap, their use was examined in more detail:

a) Chordata (109 articles)

Topics discussed in references allocated this code included:

TABLE D. GENERAL BIOSYSTEMATIC CODE CATEGORIES

General BioSystematic code	BA Previews Issue	Total no. of refs retrieved	No. of refs not mentioning organisms by name	No. of refs covering general principles but also including organisms	No. of refs dealing specifically with one or a few organisms
01000 (Micro-organisms)	51/23	86	70	7	9
85000 (Chordata)	51/20	109	85	12	12
85150 (Vertebrata)	51/20	48	31	10	7
TOTALS		243	186	29	28

(i) theoretical concepts e.g.	mathematical models, cybernetics, philosophical discourses,
(ii) general biology e.g.	conditioning theory, general dietary needs, disease epidemiology,
(iii) general methodology e.g.	specimen preparation, body fluid preservation, histological techniques, blood flow measurements, chemical levels in blood,
(iv) general biochemistry e.g.	biosynthesis, enzyme activity,
(v) general medicine e.g.	diagnostic techniques.

b) Vertebrata (48 articles)

Topics discussed in references allocated this code included:

(i) general textbooks,	
(ii) general fauna e.g.	wild life conservation, evolutionary studies,
(iii) general methodology e.g.	X-rays, mathematical calculations, anatomy and histology, serological techniques,
(iv) general biology e.g.	enzyme activity, immunity, nutrition, disease, pharmacological testing,

Our conclusions regarding the use of codes for Chordata and Vertebrata are as follows:

(a) No clear distinction could be made between the general topics discussed in either group.

(b) Topics were applicable to a wide range of animals; the titles or abstracts frequently contained terms such as: 'domestic' or 'laboratory' animals, 'fauna', 'wildlife', 'living organisms', 'vertebrates', without further specification.

(c) An article was allocated either the code for Vertebrata, or the code for Chordata, but never both. Furthermore, the article was never allotted another of the general BioSystematic codes e.g. SYS85700 (Mammalia).

(d) Of the 157 references assigned the code for either Chordata or Vertebrata, 11 (7%) named particular animals either in the title or among the enrichment terms and were also indexed with the appropriate, more specific, BioSystematic code. For example, the following articles were assigned both a general and a specific code:

"Studies on deep-freezing of boar semen. Part III. Additional effects of hyaluronidase on livability of boar spermatozoa"
(51, No.20, Abs.112556)

"Uterine weight and pituitary luteinizing hormone content in the pseudopregnant hamster"
(51, No.20, Abs.113017)

"Regeneration of the African Long Fish Proopterus"
(51, No.20, Abs.111750)

(e) A further eight references (5%) appeared to deal with clinical studies which should have been assigned the code SYS86215 (Hominidae). These references, which would have been missed by a profile designed to retrieve aspects of human medicine, were:

abstract 114036 - enriched with 'human', but SYS86215 not allocated

abstract 113339 - dealt with clinical experiments

abstract 114121 - dealt with clinical aspects of tumors

abstract 114120 - dealt with malignant tumors in the USSR

abstract 114273 - dealt with leukemia cases

abstract 111952 - dealt with the clinical use of the fibroscope

abstract 113272 - dealt with heart disease

abstract 113930 - dealt with drug poisoning

5. Overlap between CROSS and BioSystematic codes

There are several code headings which appear in both CROSS and the BioSystematic indexes. For example:

a) Micro-organisms: CRS29500 and SYS01000

b) Paleontology

Paleobotany	:	CRS50000 and SYS29000
Paleobiology	:	CRS64500 and SYS27000
Paleozoology	:	CRS63000 and SYS31000

It is understood that current BIOSIS editorial policy is that CROSS codes are assigned on the basis of the subject matter of the paper, and the BioSystematic codes are allocated according to the organisms mentioned.

a) Micro-organisms

A profile comprising the CROSS code CRS29500 and the BioSystematic code SYS01000 was run against BA Previews 51(23). Of the 91 references retrieved, it was found that:

63 articles were coded with SYS01000 only

28 articles were coded with CRS29500 of which 23 were also coded with SYS01000.

It appeared that the CROSS code was used where the paper discussed a fundamental activity of micro-organisms, e.g. enzyme formation by bacteria, microbial action in cereal digestion. In the 23 articles which had been given the BioSystematic as well as the CROSS code, named micro-organisms were mentioned in the enriched title.

The BioSystematic code SYS01000 was used to indicate that micro-organisms (usually unspecified) were relevant to some aspect of the article. It was applied to articles in which general terms such as 'microbial' or 'microflora' were used or in which the reference to micro-organisms was inferred. It was also used in conjunction with BioSystematic codes for named bacteria, fungi, etc, usually when these were being used to demonstrate a general principle.

b) Paleontology

A profile consisting of the three CROSS codes and the three BioSystematic codes concerned with paleontology was run against two issues of BA Previews, i.e. 51(24) and 52(02). The profile retrieved 137 articles which had been assigned one or more of the paleontological codes. In 121 of these articles it appeared that these codes were allocated according to the following principles:

(i) the appropriate BioSystematic code was allocated whenever an organism was mentioned in the title, enrichment terms or abstract of a paper dealing with a paleontological topic. Specifically:

SYS31000(Paleozoology) was assigned to 113 papers referring to fossil animals

SYS29000(Paleobotany) was assigned to 20 papers referring to fossil plants

SYS27000(Paleobiology) was assigned to only four papers.

(ii) the appropriate BioSystematic code was accompanied by the equivalent CROSS code, i.e.

SYS31000(Paleozoology) was assigned with CRS63000(Paleozoology)

SYS29000(Paleobotany) was assigned with CRS50000(Paleobotany)

(iii) the only paleontological CROSS code to appear without the corresponding BioSystematic code was CRS64500 (Paleobiology) which was used for general paleontological principles e.g. stratigraphy, analytical techniques.

Of the 16 articles (out of 137) which did not conform to these principles:

(i) 3 had no indication in either title or abstract as to why SYS29000 or SYS31000 had been added (the original paper was not examineed),

(ii) 10 had no CROSS code corresponding to the BioSystematic codes which had been allocated,

(iii) 3 had no paleontological BioSystematic code corresponding to the CROSS code. These were:

"The skin of Egyptian mummies. A study in survival" = histological techniques climate
(51, No.24, Abs.134412)

"Planktonic foraminifera differential production and expatriation off Baja, California" = Mexico kummerform fossils
(51, No.24, Abs.133307)

"Remarks on the genera Eucalycoceras and Protocanthoceras ammonoidea" = Pseudocalycoceras new genus fossil

(51, No.24, Abs.133364)

C. RECOMMENDATIONS FOR USING BIOSYSTEMATIC CODES

There are 18 codes describing categories of micro-organisms,
393 codes describing categories of plants,
325 codes describing animals (including 26 extinct categories),
3 codes describing categories of paleontology,
2 codes describing categories of pesticides.

The studies described above indicate that BioSystematic codes are allocated with a high degree of consistency to any specific organism or group of organisms which is mentioned in the enriched title or the abstract of the paper. In the construction of a search profile, the decision whether to use the specific name or the BioSystematic code as a search term will depend on the number of organisms which are of interest, and the amount of information required about them.

1. Phylogenetic considerations

a) BioSystematic codes are more suitable for the retrieval of broad categories of organisms such as:

Phyla,	e.g. Protozoa
Classes	e.g. Mammalia
Orders,	e.g. Marsupiala
Families,	e.g. Canidae

b) If only one or a few genera or species are required, it is preferable to search for the generic or common names rather than for the BioSystematic code, which would retrieve much unwanted material. For example, BA Previews 51(12) includes 48 articles coded SYS85306 (Anura) which covers all toads and frogs. Only 13 of these articles refer to the genus Rana (frog) but there is no BioSystematic code specific to this genus. Thus, to retrieve only those articles which refer to Rana, a word search is preferable.

c) When only Homo sapiens is of interest, the code SYS86215 (Hominidae) should be used rather than a list of all the possible synonyms for humans.

2. Subject range of BioSystematic codes

The range of topics covered by articles which were coded with selected BioSystematic codes are summarised below.

a) Codes for Viruses

These were applied to papers dealing with:

classification of viruses
disease etiology
culture growths
biochemical reactions and metabolism
antiviral agents
carcinogenesis
diagnostic pathology
DNA replication and transcription
biophysical techniques for identification

b) Codes for Coniferopsida

These retrieved papers dealing with:

systematics and classification
drugs derived from conifers
palynology
forest ecology
wood products e.g. sawdust
plant diseases affecting conifers
allergies caused by conifers
their use as animal food

c) Codes for Echinodermata

As well as papers concerning systematics, physiology, etc, these codes also retrieved articles concerning:

embryology using sea urchin eggs
new species and locations
population explosions (Crown-of-thorns)
fossil finds and their classification
their use as food (beche-de-mer)

d) Codes for Suidae

Papers retrieved dealt with:

experimental surgery
substances derived from porcine tissue
porcine cells in tissue culture
enzyme chemistry
testing of food supplements
germ-free animals
quality of pork and bacon

e) Codes for Vertebrata and Chordata

A general BioSystematic code (vertebrata, chordata, micro-organisms), but no specific code, is often assigned to papers concerned with general methods, techniques and theoretical models. Searches for these general topics should not therefore be restricted to retrieving articles which have been assigned the BioSystematic code for a specific organism. In particular some articles concerning general applications of clinical medicine are assigned the general code (vertebrata or chordata) but not the code for humans (SYS86215).

Appendix II 3

SPECIMEN PROFILE AND PRINTOUT OF REFERENCES FROM BA PREVIEWS SEARCH

1. SPECIMEN PROFILE

Profile No. W65

Statement of interest: Fate of fertilizer nitrogen applied to soil, with special reference to leaching and denitrification, and methods of nitrogen analysis.

A	W6501A0200	BREMNER J M	Author terms
	W6501A0200	CHATT J	
	W6501A0200	KOLENBRADER G J	
B	W6501TO160	NITROGEN-15	Word terms
C	W6501TO160	NATURAL ABUNDANCE	
D	W6501TWD/150	*NITRIFICATION	
E	W6501MA	CRS52805 [Soils: Physics and Chemistry]	CROSS codes
	W6501MA	CRS52807 [Soils: Fertility studies]	
F	W6501TWS/125	NITROGEN*	
G	W6501TA	FERTILIZ*	
	W6501TA	FIXATION	
	W6501TA	LEACH*	
H	W6501TWS/100	SOIL*	
I	W6501TA	NITROUS OXIDE	
	W6501TA	OXYGEN	
	W6501TA	CARBON DIOXIDE	
	W6501TA	NITROGEN*	
	W6501TA	CARBON DI OXIDE	
J	W6501TWD/75	*AEROBIC*	
K	W6501MA	CRS40000 [Soil Microbiology]	
L	W6501MWD/50	MAJ52805 [Soils: Physics and Chemistry]	Primary CROSS code
	W6501MWD/50	SUB52805 [Soils: Physics and Chemistry]	Secondary CROSS code
M	W6501MA	CRS10612 [External effects: physical and mechanical]	
N	W6501TWS/25	NITROGEN*	
O	W6501TA	PURIF*	
	W6501TA	EXTRACT*	
	W6501TA	MEASUR*	
	W6501TA	SAMPLING	
P	W6502MN	MAJ525 [Agronomy]	

Note

The left-hand column defines the logical relationships between terms. Thus, if we designate groups of terms as A,B,C ... as shown, then the logic required by the profile for retrieval of a reference is:-

A or B or C or (D and E) or (F and G) or (H and I) or (J and K) or (L and M) or (N and O) but not P.

The descending term weights are used to order the output.

The "NOT"logic" ensures that a reference allocated the printing assignment for Agronomy (MAJ525) will not be retrieved.

2. SAMPLE OF REFERENCE PRINTOUT

BEST COPY AVAILABLE

```
        W65

        BA VOL 52  NO. 22    DATE  24/01/71


ITEMS PRINTED               15
ITEMS REMAINING             999984


GIORDANO C, DE PASCALE C, ESPOSITO R,
BALESTRIERI C, CITTADINI D.
 THE DISTRIBUTION OF NITROGEN-15 IN THE AMINO-
 ACIDS OF HEMO GLOBIN FROM NORMAL SUBJECTS AND
 THOSE WITH UREMIA TREATED WITH NITROGEN-15 UREA
BIOCHIM APPL
15 (6). 1968 382-386.
ARTICLE WEIGHT  180              NUMBER    123028
NITROGEN-15.


AGARWAL A S, SINGH B R, KANEHIRO Y.
 IONIC EFFECT OF SALTS ON MINERAL NITROGEN
 RELEASE IN AN ALLOPHANIC SOIL=USA ORGANO
 INORGANIC COMPLEX SULFATES CHLORIDES
 NITRIFICATION ELECTROLYTE MICROBIAL
 RESPIRATION FERTILIZERS
SOIL SCI SOC AM PROC
35 (3). 1971 454-457.
ARTICLE WEIGHT  150              NUMBER    127986
NITRIFICATION, NITROGEN, SOIL.

DUDOREV M A, CHOBIT KO G L.
 APPLICATION OF BASIC FERTILIZER IN RAISING
 PINUS-SYLVESTRIS-G SEEDLINGS=NITROGEN
 PHOSPHORUS ORGANO MINERAL NUTRITION HUMUS
BYULL VSES NAUCHNO-ISSLED INST AGROLESMELIOR
5 (57). 1969 TRANS. 40-43.
ARTICLE WEIGHT  125              NUMBER    124709
NITROGEN.

LANG N J, FAY P.
 THE HETEROCYSTS OF BLUE-GREEN ALGAE PART 2
 DETAILS OF ULTRASTRUCTURE=ANABAENA-CYLINDRICA
 PORE CHANNEL MICRO PLASMODESMATA PLASMALEMMAS
 THYLAKOIDS MEMBRANES NITROGEN FIXATION
PROC R SOC LOND B BIOL SCI
178 (1051). 1971 193-203.
ARTICLE WEIGHT  125              NUMBER    125866
NITROGEN.

WANEK W, ONDRACEK L, HAMPL J.
 COVALENT NITROGEN PHOSPHORUS COMPOUNDS AS
 SOURCE OF PLANT NUTRITION=OATS-M FERTILIZER
 CHEMICAL PHYSICAL PROPERTIES
Z PFLANZENERNAEHR BODENKD
128 (3). 1971 169-180.
ARTICLE WEIGHT  125              NUMBER    127917
NITROGEN.
```

3. TERM FREQUENCY LISTING

```
SEARCH TERM FREQUENCY

PROFILE NO   V65-Y          RUN AGAINST   BIOLOGICAL ABSTRACTS PREVIEWS    VOL  52  ISS  22        21/01/71

             SEARCH TERM                                                     FREQUENCY

        NITROGEN*           .     .     .     .     .     .     9
        SOIL*               .     .     .     .     .     .     6
       *NITRIFICATION       .     .     .     .     .     .     2
        NITROGEN-15         .     .     .     .     .     .     1
```

BEST COPY AVAILABLE

Appendix II 4

Experimental Information Unit
7 Keble Road, Oxford

EXPLANATION OF BIOLOGICAL ABSTRACTS OUTPUT (1970)

The first sheet of your output consists of your Profile Printout, showing the logical sequence of terms and codes, grouped in parameters, and indicating the weight assigned to each. For subsequent computer runs of this profile, you will not receive a profile printout but if your profile is amended, i.e. a new profile is established, you will receive a printout of this the first time it is run.

The second sheet shows the Search Term Frequency, i.e. the number of times one of the terms or codes in your profile produced a 'hit'.

The remaining sheets are your Printout of references for one issue of Biological Abstracts or BioResearch Index.

Example: On the specimen Printout attached, the following information is given:-

BA Vol.51 No.03 Date 15/03/70
(This is the printout for Biological Abstracts, Volume 51, Issue No. 3, which was run on 15 March 1970)

Items printed: 16
(16 references were retrieved)

Journal reference (example)

TKOCZ, C, KUHN, K (1)

The formation of triple helical collagen molecules from alpha 1 or alpha 2 polypeptide chains (2) = calf rat skin (3)

Eur J Biochem (4)

7$^{(5)}$ 4$^{(6)}$ 1969$^{(7)}$ 454-462$^{(8)}$

Article weight 100 (9)

Collagen$^{(10)}$ CRS10064 CRS10506$^{(11)}$ Number 12363$^{(12)}$

(1) Author(s)
(2) Article title
(3) 'Enrichment' words added to title by BA editors

Explanation of Biological Abstracts Output

(4) Abbreviated journal title

(5) Volume

(6) Issue

(7) Year

(8) Pages

(9) Weight of this article, i.e. sum of weights of terms and codes which have produced a 'hit'

(10) Term in profile which has produced a 'hit'

(11) CROSScode in profile which has produced a 'hit'

(12) Abstract Number in Biological Abstracts or Citation in BioResearch Index

CROSScodes and BioSystematic Codes

The CROSScode numbers correspond to the subject area headings in Biological Abstracts. The paper is cross-indexed and it is possible to distinguish between three types of CROSScodes:-

CRS: Subjects in the CROSScode list, representing all subject areas covered by the article. The codes which appear in the printed CROSS-Index in Biological Abstracts are differentiated as:

MAJ: The principle subject area, i.e. that under which the abstract can be found in the printed Biological Abstracts.

SUB: Important subject areas to which the article is cross referenced.

N.B. If MAJ and SUB are not specified and only CRS appears in the profile, all three types will be retrieved.

September 1970

APPENDIX III

CONTENTS

1. TABLE: LIST OF ESTABLISHMENTS. MARCH, 1971.

2. "BA PREVIEWS EXPERIMENT IN THE U.K." [HANDOUT TO USERS]. APRIL, 1970.

3. QUESTIONNAIRE. [LITERATURE SEARCHING HABITS].

4. TABLE: GROUP AND SHARED PROFILES.

5. [FEEDBACK CARD].

6. TABLE: ANALYSIS OF COMMENTS ON FEEDBACK CARDS RETURNED BY A SAMPLE OF 100 USERS.

7. [LETTER TO ALL USERS WITH EXPLANATION]: EVALUATION OF THE BA PREVIEWS MAGNETIC TAPE SERVICE. MAY, 1971.

8. TABLE: EVALUEES BY CATEGORY OF EMPLOYMENT.

9. TABLE: EVALUEES BY SPECIFIC SUBJECT CATEGORY.

10. TABLE: MANUAL LITERATURE SEARCHING: PRIMARY JOURNAL USAGE. LIST OF PRIMARY JOURNALS CITED BY 10 OR MORE USERS.

11. [LETTER TO USERS WHO FAILED TO RETURN FEEDBACK]. FEBRUARY, 1971.

12. [LETTER TO DELETED USERS AND] QUESTIONNAIRE: REASONS FOR WITHDRAWAL FROM THE BA PREVIEWS EXPERIMENT. JUNE, 1971.

13. A SAMPLE OF USER REQUIREMENTS.

Appendix III 1

TABLE: LIST OF ESTABLISHMENTS

Type	Establishment	Location	No. of Profiles	
A. Academic	1. Oxford University	Oxford		
	SCIENCE: Agriculture		2	65
	Biochemistry		8	
	Biomathematics		2	
	Botany		5	
	Forestry		2	
	Genetics		3	
	Geology		2	
	Physiology		8	
	Zoology		13	
	MEDICINE: Human Development		2	
	Obstetrics and Gynaecology		12	
	Ophthalmology		5	
	Pathology		1	
	2. Nottingham University	Nottingham		
	SCIENCE: Botany		17	62
	Pharmacy		5	
	Zoology		20	
	AGRICULTURE: Agricultural Economics	Sutton Bonnington	2	
	Applied Biochemistry		7	
	Physiology		11	
	3. Queen's University of Belfast	Belfast		
	Biochemistry		1	2
	Medical Biology		1	
	4. Cambridge University	Cambridge		
	Geology		2	3
	Physiology		1	
	5. Dublin University College	Dublin		
	Botany		1	2
	Medicine		1	
	6. East Anglia University	Norwich		
	Biological Sciences			1
	7. Hatfield Polytechnic	Hatfield		
	Biological Sciences			5
	8. Liverpool University	Liverpool		
	Botany			3

Type	Establishment	Location	No. of Profiles
A. Academic (contd)	9. London University, Middlesex Hospital Medical School	London	
	Biology		1
	10. Loughborough University	Loughborough	
	Chemistry		1 } 5
	Ergonomics		4
	11. Manchester University	Manchester	
	Biological Chemistry		2
	12. Sheffield University	Sheffield	
	Anatomy		2
	Pharmacology		2 } 8
	Physiology		4
	13. University of Wales, Welsh National School of Medicine	Cardiff	6
B. Independent Research Establishments	1. Animal Health Trust	Newmarket	3
	2. Marine Biological Association of UK	Plymouth	9
	3. Royal Society	London	1
C. Government			
a) Home Office	Central Research Establishment	Aldermaston	2
b) Min. of Agric. Fisheries and Food	1. Central Veterinary Laboratory	Weybridge	10
	2. National Agricultural Advisory Service	Cambridge	2
		Reading	2
		Derby	14
	3. Plant Pathology Laboratory	Harpenden	3
	4. Royal Botanic Gardens	Kew	1
	5. Sea Fisheries Research Laboratories	Lowestoft	2
c) Min. of Technology	Torry Research Station	Aberdeen	3
d) Agric. Research Council	1. Agricultural and Horticultural Research Station	Long Ashton	6

Type	Establishment	Location	No. of Profiles
C. Government (contd)			
d) Agric. Research Council	2. Institute for Research on Animal Diseases	Newbury	12
	3. Meat Research Institute	Bristol	5
	4. National Institute for Research in Dairying	Reading	9
	5. National Vegetable Research Station	Wellesbourne	18
	6. Radiobiological Laboratory (Letcombe)	Wantage	7
	7. Rothamsted Experimental Station	Harpenden	11
	8. Unit of Muscle Mechanism and Insect Physiology	Oxford	1
	9. Weed Research Organisation	Oxford	4
e) Medical Research Council	1. Radiobiology Unit	Harwell	4
	2. Unit for Metabolic Studies in Psychiatry	Sheffield	2
f) Natural Environment Research Council	1. Fresh-Water Biological Assoc. of the U.K.	Windermere	1
	2. Nature Conservancy	Abbots-Ripton	11
g) EIRE, Dept. of Agric. and Fisheries	Fisheries Division	Dublin	1
h) EIRE, Institute for Industrial Research and Standards	Industrial Research Centre	Dublin	1
i) General	1. Trent River Authority	Derby	5
	2. British Museum (Natural History)	London	5
D. Industry	1. Boots Pure Drug Co. Ltd.	Nottingham	5
	2. Fisons Pharmaceuticals	Loughborough	20
	3. Imperial Metal Industries Ltd.	Birmingham	1
	4. May and Baker Ltd.	Dagenham	3

of these articles it appeared that these codes were allocated according

---

Type	Establishment	Location	No. of Profiles
D. Industry (contd)	5. Plant Protection Ltd. (ICI)	Bracknell	5
	6. Lord Rank Research Centre	High Wycombe	6
	7. Wellcome Research Laboratories	Beckenham	11
	8. Harry Wheatcroft and Sons Ltd.	Nottingham	1
F. Naturalists	Mr. Davies	Nottingham	1
	Dr. Kent	Nottingham	1
	Mr. Sawyer	Newport	1
	Dr. Veal	Nottingham	1
G. Information Centres	1. Bee Research Assoc.	Gerrards Cross	1
	2. Biodeterioration Centre	Birmingham	1
	3. Biomedical Information Centre	Sheffield	3
	4. Commonwealth Bureau of Animal Health	Weybridge	1
	5. Commonwealth Bureau of Helminthology	St. Albans	3
	6. Commonwealth Forestry Bureau	Oxford	1
	7. National Documentation Centre for Sport, P.E. and Recreation	Belfast	2
	8. Project FAIR (MRC)	London	4
	Total no. of Profiles	=	392

Appendix III 2

Experimental Information Unit
7 Keble Road, Oxford

BA PREVIEWS EXPERIMENT IN THE UK

1. Introduction

BA Previews consists of a magnetic tape for each semi-monthly issue of Biological Abstracts and each monthly issue of BioResearch Index, i.e. a total of 36 tapes per year.

With the exception of the actual text of the abstract, all the information contained in the main text and the indexes of the printed Biological Abstracts and BioResearch Index is now available on tape. The tapes are distributed approximately one month earlier than the corresponding printed issues of the journal.

The annual lease of the BA tapes is $3800, plus the cost of the tapes.

2. Purpose of the Experiment

The Office for Scientific and Technical Information (OSTI), a section of the Department of Education and Science, has been offered a year's free supply of BA Previews tapes for use in the UK and has awarded grants of £14,048 and £5650 to the Universities of Oxford and Nottingham for a two-year investigation of the use of the tapes.

The results should provide a basis for deciding whether a service based on BA Previews should be set up for users of biological information in this country. In this context, the following aspects of the service will be investigated:-

(a) Journal coverage, degree of currency, logical organization of the indexes and the allocation of index words and codes.

(b) The best way of formulating computer-usable profiles representing user interests.

(c) The most acceptable form for a BA Previews service, i.e. Current Awareness only, or Current Awareness and Retrospective.

(d) The relative performance of a BA Previews service compared with other mechanized and conventional services covering the same fields.

In addition, it is hoped to obtain some general information on the pattern of information needs and usage among workers in biology and the subject fields covered by BA Previews.

3. Staffing

Three liaison scientists have been appointed, two at Oxford and one at Nottingham.

OXFORD

Experimental Information Unit
7 Keble Road, Oxford OX1 3QL

Tel: Oxford 52427

1. Miss Margaret Williams, B.Sc.,Dip.Lib.,A.L.A a Botany and Zoology graduate from the University of Leeds, formerly Division Librarian at Imperial Chemical Industries Limited, Pharmaceuticals Division, and with varied experience in industrial and academic library and information work.

2. Miss Barbara Smith, M.Sc., a Zoology graduate from the University of Melbourne with research experience in vertebrate reproductive physiology, and formerly medical abstracter at Derwent Publications Ltd.

NOTTINGHAM

Department of Botany
University of Nottingham
Nottingham NG7 2RD

Tel: 56101, extn. 3143

Mrs Janet Stow, M.Sc., a sandwich course graduate in Applied Biology from Salford University, with research experience in cell culture.

Their task is to build up a limited population of users, who will receive regularly a printout selected by the computer to meet their individual interests. The service is provided free in exchange for feedback from the users about its performance.

Within restricted areas around Oxford and Nottingham, it is hoped to interview and offer participation to all potential users of the literature covered by BA Previews. This includes scientific papers, books and conference proceedings in a broad area of biology, as well as non-clinical medicine, agricultural sciences, biochemistry and pharmacology. It is proposed to recruit not only university staff and research students, but also technical college staff, workers at government and industrial research stations and in industry and some school-teachers and amateur naturalists.

4. The role of the user in the experiment

(a) First interview

If you would like to participate in the experiment you will be asked to have an interview (lasting approximately 1½ hours) with one of the liaison scientists. This will commence with the completion of a short questionnaire, which will help us to ascertain your background and literature searching practice. We shall then ask you for a statement of your research interests, from which we will proceed to compile a 'user profile'.

The user profile consists of a logical co-ordination of words and any of the following categories of codes:-

(i) Single or compound words: Appearing in titles of relevant articles,
Added by BA to enrich titles

(ii) CROSS Index Codes

(iii) BioSystematic Index Codes

(iv) Authors

(v) Journal Codens

In the computer searching, the terms (or codes) in the profile are matched against the record on the BA Previews tape and a printout of relevant references is produced from the 'hits'.

Each month you will receive three printouts, two corresponding to the fortnightly issues of Biological Abstracts and one to the monthly issue of BioResearch Index. Accompanying each printout will be a small feedback card, on which you are asked to mark which references are Important or Relevant and return it to the liaison scientist. The following definitions should assist you in making a decision:-

IMPORTANT: a reference directly related to your research interests which you will read as soon as possible, i.e. it is essential that you know of it.

RELEVANT: a reference which though not important, is of some interest to you, i.e. you are glad that you were notified of it.

The important category should be interpreted restrictively and the relevant category generously.

You may already know of (or have read) a reference before it appears in the output. However, your relevance judgments should be independent of whether a reference is familiar to you or is 'out-of-date'.

(b) Second interview

When you have received a few initial printouts, you will be contacted by the liaison scientist for a second interview (approximately ½ hour) to discuss the effectiveness of the search, so that any necessary amendments to your profile can be made. If, however, you consider the search to be of little value to your work, it would be advisable for you to discontinue as a user.

It is hoped that the requisite number of participants will have been acquired by the late summer, so that the experiment can become fully operational. It will then run for one year, at the end of which there will be a final interview with the liaison scientist to assess the value of the experiment to the user.

[The specimen profile and the annotated sample of computer printout, which were distributed to users with the above circular, are not included in this appendix. Samples of both a profile and a printout appear in Appendix II. 3].

Appendix III 3

BA PREVIEWS QUESTIONNAIRE

1. Name: Profile No:

2. Address:

3. Type of organisation
   - (a) University or College (incl. Teaching Hospital)
   - (b) Independent Research Establishment
   - (c) Government Establishment
   - (d) Industry
   - (e) School
   - (f) Free-lance naturalist

4. No. of years research experience

5. Background

   Are you scanning/searching exclusively for your own work? Yes/No

   If No, please expand:

6. Manual literature searching practice

   (a) Do you see certain primary journals regularly? Yes/No

   List the 10 most important primary journals (in order of relative importance)

   (i)
   (ii)
   (iii)
   (iv)
   (v)
   (vi)
   (vii)
   (viii)
   (ix)
   (x)

(b) Of the following list of secondary journals, which, if any, do you search regularly:

Title	Regularity		Availability
	C	R	
Biological Abstracts			
Chemical Abstracts			
Chemical Titles			
C.B.A.C.			
Current Contents			
Index Medicus			
Internat. Abs. Biol. Sci.			
Bulletin Signaletique			
Excerpta Medica			
British Medical Abstracts			
Science Citation Index			
Zoological Record			

Regularity: C = Current awareness purposes
R = Retrospective Searches
✓ = Prime Tool
X = Occasional Tool

Availability: L = Own Library
Circ. = Circulation
B = Nearby building
D = > 10 minutes away

(c) Do you use any specialised secondary journals (e.g. Microbiol. Abs.) including Indexes, Bulletins Yes/No

If yes, specify:

(d) Is your searching adequate Yes/No

If no, give reasons:

7. Mechanised Services

(a) Have you used any of the following mechanised services: Yes/No

(i) Medlars - Monthly
Retrospective

(ii) ASCA

(iii) CT
CBAC
POST-J

(iv) Ringdoc

(b) If yes, have you an existing profile? Yes/No

(c) If yes, has this eliminated the need for manual searches of the secondary journals? Yes/No

(d) If yes, has this reduced the number of primary journals you now scan? Yes/No

If no, to all,
Are you familiar with an IR computer printout? Yes/No

8. Use of Biological Abstracts

Do you use Biological Abstracts? Yes/No

If yes,

(a) Do you use BA for

(i) Current Awareness	Regularly/Spasmodically/Never
(ii) Retrospective searches	Frequently/Infrequently/Never

(b) Indexes: do you consult:

	C	R
AUTHOR		
B.A.S.I.C.		
Cross-code		
Taxonomic code		
Appropriate subject group		

Comments on design and presentation:

(c) Do you think the journal coverage is adequate? Yes/No

If no, what titles would you like to see included?

(d) Do you consider the abstracts are sufficiently informative? Yes/No

(e) Does the time-lag between publication of an article in a primary journal and its appearance in BA render the latter inadequate as a Current Awareness Tool for you? Yes/No

If no,

Is your non-use of BA attributable to:

- (i) No subscription to it by organisation
- (ii) Inaccessibility
- (iii) Do not like format of indexes
- (iv) Do not consider journal coverage adequate for own subject field
- (v) Rely on another service
  i.e.
- (vi) Do not find any need for abstract services
- (vii) Tend to rely on "Invisible Colleges", i.e. personal contacts with co-workers in own field.
- (viii) Other, specify:

Appendix III 4

TABLE: GROUP AND SHARED PROFILES

Total no. of group and shared profiles written	47
No. of group and shared profiles deleted June 1971	5
No. of group profiles participating in the evaluation	20
No. of shared profiles i.e. only one person was interviewed, acting on behalf of a group	12
No. of group profiles i.e. more than one person was interviewed	35

	Group of 2	Group of 3	Group of 4	Group of > 4	Total
Evaluees	13	4	1	-	18
Non-evaluees	8	3	2	1	14
Deleted profiles	1	1	1		3
Total	22	8	4	1	35

Appendix III 5

FEEDBACK CARD
(before October 1970)

BA - PREVIEWS

NAME: ____________________

PROFILE NO.________ ISSUE NO.________

References retrieved: Total No.____________

Important____________

Relevant ____________

Comments:

FEEDBACK CARD

(after October 1970)

BA - PREVIEWS

NAME: ______________________________

PROFILE NO. ________ ISSUE NO. ________

References retrieved: Total No. ______________

Major Value ______________

Minor Value ______________

Comments:

Appendix III 6

TABLE: ANALYSIS OF COMMENTS MADE ON FEEDBACK CARDS RETURNED BY A SAMPLE OF 100 USERS

Comments	No of Users making this comment
A. PROFILE CONSTRUCTION	
(a) Retrieval of unwanted terms, codes or journals (specified): suggestions for removal	26
(b) Retrieval of unwanted subjects (unspecified)	22
(c) Subsequent to amendment of profile	
(i) General satisfaction	18
(ii) General dissatisfaction	4
(iii) No difference	1
(d) Subject interests _not_ adequately satisfied	13
(e) Subject interests adequately satisfied	10
(f) Truncation of words resulting in retrieval of unwanted references: suggestions for removal	10
(g) Spelling error in profile (specified)	8
(h) Additional terms (specified) for inclusion in profile: suggestions	6
(i) Lack of understanding of coding of profile	5
(j) Retrieval of too many references: _no_ suggestions for decreasing	4
(k) Retrieval of specified unwanted authors, e.g. BROWN R retrieves BROWN RB	2
B. VALUE OF REFERENCES	
(a) References from journals not scanned by user	10
(b) General satisfaction with printout	8
(c) Most of the MAJOR VALUE references already seen	8
(d) Some of the references are very old	5
(e) MAJOR VALUE references retrieved which had been missed by the user in his own searching	4
(f) Reference to a paper written by the user	4

Comments	No. of Users making this comment
(g) References not available from NLL	3
(h) References inaccessible, but interesting	2
(i) Query: Why was a certain reference retrieved?	2
C. RELEVANCE DECISIONS	
(a) High proportion of relevant articles on printout, as most of them were from one highly relevant publication, e.g. Symposium, Conference proceedings	8
(b) Difficult to distinguish between MAJOR VALUE and MINOR VALUE	4
(c) Confusion between IMPORTANT and RELEVANT (i.e. IMPORTANT included with RELEVANT references)[2]	3
(d) Some of the irrelevant references are of 'fringe' interest: would like a third relevance category	2
(e) Difficult to determine relevance from title only	2
D. FORMAT OF PRINTOUT	
(a) Difficult to read	16
(b) Misunderstanding of bibliographic details (e.g. Issue numbers of BA and BRI are in different sequences)[3]	6
(c) Easier to read than usual, i.e. better quality printout	6
(d) Lack of understanding of BIOSIS notation (i.e. the addition of -D and -M to plant names to indicate di- or mono- cotyledons)	2
(e) Part of bibliographic citation missing: full title requested	2
(f) Search term frequency list missing	2
(g) Dislike of green-lined paper	1
(h) Printout sheets give pagination but not profile no. or issue no., so could become misfiled	1

Foot Notes to Table

[1] This figure refers to the number of individual users who made this comment on a feedback card, irrespective of the number of times they made it.

[2] This confusion was eliminated by the replacement of IMPORTANT and RELEVANT with two mutually exclusive categories, i.e MAJOR VALUE and MINOR VALUE.

[3] To assist users in understanding the printout, the leaflet "Explanation of BIOLOGICAL ABSTRACTS' output", which is shown in Appendix II.4, was distributed.

Appendix III 7

BA—PREVIEWS PROJECT

UNIVERSITY OF OXFORD
Miss Margaret I. Williams, B.Sc., Dip.Lib., A.L.A.
Miss Barbara Smith, M.Sc.

UNIVERSITY OF NOTTINGHAM
Mrs. Janet Stow, M.Sc.
(Department of Botany)

EXPERIMENTAL INFORMATION UNIT
7 KEBLE ROAD
OXFORD, OX1 3QL
Telephone: Oxford 52427

May 1971

In my letter of 16 April, I mentioned that I should be writing later to enquire if you would be willing to take part in the detailed evaluation phase of our experiment

The object of the evaluation is to investigate characteristics of the BA Previews service such as coverage of the literature, the currency, and the quality of indexing.

In order to assist you in deciding whether or not you are prepared to undertake this work, I am enclosing a form which gives details of the experiment and the work required from participants. If there are any points about which you are not clear, I hope you will not hesitate to contact me.

Would you kindly complete the enclosed card as soon as possible, indicating whether or not you are willing to take part as an evaluee.

Even if you are not able to act as an evaluee, we hope you will continue to return the feedback cards. I am sure you will appreciate that it is expensive to run an experiemental computerised information service such as BA Previews. If I do not hear from you by 31 May 1971, I shall assume that you no longer wish to receive printouts and your profile will be deleted.

Yours sincerely,

Margaret Williams
Senior Liaison Scientist

BV-4

Experimental Information Unit
7 Keble Road, Oxford OX1 3QL

EVALUATION OF THE BA-PREVIEWS MAGNETIC TAPE SERVICE

Evaluation of the BA-Previews service will take into account the following factors:

1. The proportion of references retrieved which are relevant
2. The proportion of references 'missed' by the service
3. The cost of providing the service
4. Acceptability to the user.

A detailed analysis will be made of relevant references which were not retrieved or irrelevant references which were retrieved. It is hoped that this 'failure analysis' will indicate ways of improving the existing service.

Information required from participants in our evaluation will be:

1. In order to identify the proportion of relevant references missed by the service, you will be asked to report relevant references found by your usual methods of searching the literature (e.g. Primary Journals, Abstract Journals, Titles Lists, 'Personal Communication'). This feedback will be for a maximum period of six weeks, during June - July.
2. You will also be asked to identify all relevant references appearing in 4-6 issues of the BA computer output and indicate whether you have already seen these references. This second set of feedback will be required during the period July to October. The exact period can be varied to take account of summer holidays,etc.
3. Some users will be also asked to make relevance judgments on an output from another computer service such as Chemical Titles or MEDLARS.

Special forms will be provided for listing the relevant references, and detailed instructions will be sent out at each stage of the experiment. In order to identify a reference it will only be necessary to quote first author, journal abbreviation, volume number, year and first page number.

/continued

Evaluation of the BA-Previews Magnetic Tape Service (continued)

We recognise that the provision of this detailed feedback will involve you in some extra work and in designing the experiment we have endeavoured to keep this to a minimum. The amount of work which will be required of you should occupy not more than one hour of your time in any one week. If, after agreeing to take part in the experiment, you find that the evaluation is taking up too much of your time, the period in which you send feedback can be reduced, e.g. reporting of references found by 'your usual methods' can be limited to two instead of six weeks.

At the end of the experiment all participants, including those not able to take part in the full evaluation, will be asked to complete a final questionnaire. This will provide an opportunity for you to comment on the merits of the service, and also on the methods which we have used for collecting data.

Evaluation of BA-Previews

- * I am willing to take part in the evaluation phase of the experiment ☐
- * I am not willing to take part in the evaluation phase of the experiment, but will continue to provide regular feedback ☐
- * I wish to withdraw from the experiment ☐

* Please tick in the appropriate box

Profile No:

Name:

[Card which accompanied the circular letter BV 5].

BEST COPY AVAILABLE

Appendix III 8

TABLE: EVALUEES BY CATEGORY OF EMPLOYMENT

Employment Category	No. of Evaluees	No. of Profiles
A. Academic	56	60
B. Independent Research Establishment	8	9
C. Government:		
(a) Ministries, NERC etc.	18	18
(b) Agricultural Research Council	39	40
(c) Medical Research Council	2	2
D. Industry	21	21
F. Naturalist	1	1
G. Information Centre	6	9
TOTALS	151	160

Appendix III 9

TABLE: EVALUEES BY SPECIFIC SUBJECT CATEGORY

Subject Category	No. of Evaluees	No. of Profiles
1. Agricultural Science	9	9
2. Aquatic Biology	4	4
3. Biochemistry	15	15
4. Biometrics	-	-
5. Biophysics	6	6
6. Cell Biology	7	7
7. Ecology and Conservation	6	6
8. Embryology	1	1
9. Endocrinology	5	5
10. Entomology	6	6
11. Experimental Psychology	1	2
12. Food Science	4	5
13. Forestry	3	3
14. Genetics	2	2
15. Immunology	2	3
16. Medicine	3	5
17. Microbiology	9	9
18. Nutrition	1	1
19. Palaeontology	1	1
20. Parasitology	3	3
21. Pathology/Forensic Science	4	4
22. Pharmacology/Toxicology	7	7
23. Physiology	13	17
24. Plant Pathology	8	8
25. Plant Physiology	20	20
26. Public Health	-	-
27. Systematic Botany	3	3
28. Systematic Zoology	3	3
29. Veterinary Science	3	3
30. Virology	2	2
TOTALS	151	160

Appendix III 10

TABLE: MANUAL LITERATURE SEARCHING: PRIMARY JOURNAL USAGE.
LIST OF PRIMARY JOURNALS CITED BY 10 OR MORE USERS

Journal	No. of Citations
Nature	159
Biochemical Journal	72
Biochimica et Biophysica Acta	51
Journal of Biological Chemistry	47
Plant Physiology	43
Journal of Physiology	39
Science	36
Lancet	32
Annals of Applied Biology	29
Journal of Ecology	28
British Medical Journal	27
Journal of Experimental Botany	
Journal of Applied Ecology	25
Journal of General Microbiology	
Biochemistry	23
Journal of Economic Entomology	
Proceedings of the Academy of Sciences of the United States of America	22
Archives of Biochemistry and Biophysics	21
Phytopathology	
Planta	20
New Phytologist	19
American Journal of Physiology	18
Analytical Biochemistry	
Journal of Molecular Biology	
Plant Pathology	
Veterinary Record	
Journal of Bacteriology	17
Journal of Immunology	
Journal of Marine Biology	

Journal	No. of Citations
Physiologia Plantarum	
Journal of Cell Biology	16
Journal of General Virology	
Ecology	15
Endocrinology	
Journal of Animal Ecology	
Journal of Endocrinology	
Plant and Soil	
Virology	
Biochemical and Biophysical Research Comms.	14
Canadian Journal of Botany	
Journal of Comparative Pathology	
Australian Journal of Biological Sciences	13
Journal of Cell Science	
Journal of Reproduction And Fertility	
Journal of the Science of Food and Agriculture	
Plant Disease Reporter	
American Journal of Botany	12
British Journal of Nutrition	
Bulletin of Entomological Research	
Immunology	
Journal of Agricultural and Food Chemistry	
Journal of Zoology	
Research in Veterinary Science	
Journal of Experimental Biology	11
Journal of Soil Science	
Journal of Ultrastructure Research	11
Parasitology	
Soil Science	
Weed Research	
American Journal of Obstetrics and Gynecology	10
American Journal of Veterinary Research	
Annals of Botany	
Canadian Journal of Microbiology	

BEST COPY AVAILABLE

Journal	No. of Citations
Journal of Insect Physiology	
Journal of Nutrition	
Transactions of the British Mycological Society	

Appendix III 11

BA PREVIEWS PROJECT

Experimental Information Unit,
7 Keble Road,
Oxford OX1 3QL

February 1971

Dear

When we invited you to participate in our experiment, we told you that we were offering you a free computerized information service. However, we stressed the importance to us of a regular feedback on its performance.

On checking our records, I find that I have not received a feedback card from you since the one for BA Issue No. (run on). I should be very grateful if you would kindly fill in the outstanding cards and return them to me as soon as possible, so that I can complete my records.

In March, we hope to commence the detailed evaluation phase of our experiment, when it will be essential for us to have information from our users. If I hear nothing from you before the end of February, I shall assume that the service is of no potential interest to you and you no longer wish to receive BA printouts. In this case, I shall be obliged to discontinue sending them to you.

Yours sincerely,

Appendix III 12

BA PREVIEWS PROJECT

28th June 1971

Dear

I am writing to inform you that, as you have indicated that you wish to withdraw from the experiment, your profile has now been deleted and you will receive no further print-outs.

However, I should be interested to know your reason for withdrawing. Accordingly, I am attaching a brief questionnaire, listing possible reasons and would be grateful if you would kindly mark the relevant one(s).

A stamp addressed envelope is enclosed for your reply.

Yours sincerely,

Encl.

Appendix III 12

Questionnaire

Reasons for withdrawing from the BA-Previews Experiment

1. BA-Previews Service has proved of no value :-
    - a) Have already seen most of the references
    - b) The references are too old to be of any use
    - c) The coverage of BA-Previews is inadequate for my research topic
    - d) The print-out includes too many irrelevant references
2. Rely on another service :-
    - a) Abstracting or Indexing Journal
    - b) Library/Information Service
    - c) Invisible College
    - d) Other (please specify)
3. Analysis of print-out and return of feed-back is too time-consuming to be justified
4. Format of the print-out is unacceptable (e.g. illegible, difficult to file)
5. Research topic has changed, rendering profile inadequate
6. Leaving the U.K. for work abroad
7. Other (please specify)

Please indicate with a ✓ the main reason for your withdrawing from the experiment. If there are further reasons for your dissatisfaction with the service, please indicate these with a X

Appendix III 13

A SAMPLE OF USER REQUIREMENTS

One line summaries of the statement of requirements of twenty users are given below. It must be emphasised that many requirements were more complex than can be indicated in a single sentence.

1. Mechanism of action of antimicrobial compounds.
2. Environmental physiology and psychology in humans, in particular the influence of heat, noise and light.
3. Soil physics, management, hydrology and aeration.
4. Salt and water balance in the lower vertebrates.
5. Biology of the Tettigoniidae and insect eggs.
6. Intestinal microorganisms in poultry; the effect of mycotoxins on poultry.
7. Diagnosis of tropical diseases with particular reference to parasitic diseases.
8. The digestion and absorption of proteins.
9. Pectolytic enzymes.
10. Animal waste management.
11. Ruminant nutrition and metabolism.
12. Factors affecting the growth and development of roots.
13. Biological control of insect pests in soil.
14. Developmental genetics of Drosophila melanogaster.
15. The biochemistry of the normal and cataractous lens of mammals.
16. Comparative studies of the biology of selected grasses.
17. Soil and plant factors affecting the efficiency of herbicides in weed control.
18. Preparation and use of anti-lymphocyte serum.
19. Crop diseases and their chemical control.
20. Effect of agricultural pesticides on wild life.

Appendix IV

CONTENTS

1. INSTRUCTION SHEET REGARDING PRECISION FEEDBACK: RELEVANT REFERENCES RETRIEVED BY THE BA PREVIEWS SERVICE

2. DOCUMENT SAMPLE: INSTRUCTION SHEET

3. INSTRUCTION SHEET AND FORM BV-6: REFERENCES FOUND BY SEARCHING OF CURRENT LITERATURE [Gave instructions to users about the references identified by the user in the course of his normal current awareness searches. These references formed the BV-6 sample used for computing coverage and recall ratios].

4. Instruction sheets for the alternative services which were used to provide recall and coverage samples

   CHEMICAL ABSTRACTS CONDENSATES (CAC) PRINTOUT

   MEDLARS PRINTOUT

   ASCA SERVICE [request for a list of cited authors or references]

   ASCA SERVICE: CITATION SEARCH [instructions regarding output from the citation search]

5. PROCEDURE FOR THE ANALYSIS OF COVERAGE SAMPLES

   [Method of determining whether or not a reference from the coverage sample was covered by BA or BRI]

6. TABLE: DISTRIBUTION OF APPEARANCE DATES IN BA PREVIEWS OF REFERENCES FROM COVERAGE SAMPLES

7. CURRENCY SPREAD IN BA/BRI FOR INDIVIDUAL ISSUES OF PRIMARY JOURNALS

8. TABLE: SERIAL PUBLICATIONS NOT COVERED BY BIOLOGICAL ABSTRACTS OR BIORESEARCH INDEX

Appendix IV 1

Experimental Information Unit
7 Keble Road, Oxford, OX1 3QL

BA PREVIEWS EXPERIMENT

RELEVANT REFERENCES RETRIEVED BY THE BA PREVIEWS SERVICE

INSTRUCTION SHEET

GENERAL

1. For 4 issues of BA Previews, you will receive 2 copies of your printout but no feedback cards. One copy is for your own use and the second copy for you to indicate the relevance of each reference retrieved.
2. In assessing the relevance of a reference, the original definitions, which you have been using when completing your feedback cards, apply:-
   - Major Value (MAJ): Very centrally related to your research interests and could make a significant contribution.
   - Minor Value (MIN): Of interest, and happy to have been informed of it, but of no great loss to have missed it.
   - Irrelevant (X): Of no interest

INFORMATION REQUIRED

3. Will you please mark each reference as:-
   MAJ (Major Value); MIN (Minor Value); X (Irrelevant)
4. References already known: In addition, will you please put an S against any MAJ or MIN references if you already knew of the existence of this article before you saw it on your computer printout.
5. Each reference should be marked in the outer margin, i.e. the lefthand column of references should be marked in the lefthand margin, the righthand column in the righthand margin.
6. Examples:

MAJ	Article of Major Value:	not already known
MIN/S	Article of Minor Value:	already known
X	Irrelevant	

7. The output should be returned in the enclosed addressed envelope, as soon as possible.

GROUP PROFILES

8. For most profiles, we expect only one person to be assessing relevance, even though, in some cases, several other research workers (e.g. graduate students) may see the computer printout. However, at a few interviews, more than one person attended and contributed to designing the profile and each has been making his own relevance decisions. If this is so, the following procedure should be adopted:-
   (i) Mark at the top of the output the number of people making relevance judgements (e.g. 3 people)
   (ii) Each participant should provide an independent set of decisions.

Example: If three people make relevance decisions, each reference will have 3 marks, e.g. MAJ/S
X
MIN

Appendix IV 2

Experimental Information Unit
7 Keble Road, Oxford, OX1 3QL

BA PREVIEWS PROJECT

DOCUMENT SAMPLE: INSTRUCTION SHEET

1. For the final stage of the evaluation, we are sending you a random sample of documents.

2. By comparing relevance judgements made on the BA Previews printout with those made on the actual document, we wish to determine the extent to which a research worker is able to make a correct prediction of the relevance of a document on the basis of the title alone.

3. The enclosed articles all appeared on one or other of the four duplicate printouts you received a few weeks ago and may or may not have been judged relevant by you. In this case, you were assessing relevance on the information supplied on the printout i.e. Title, Enrichment words, and Bibliographic reference.

4. After studying these articles, we should like you to indicate the relevance of each, without reference to your previous judgement.

5. Will you please mark at the top of each document:-

MAJ (Major value) / MIN (Minor value) / X (Irrelevant)

6. The documents should then be returned in the enclosed addressed envelope, as soon as possible.

Appendix IV 3

Experimental Information Unit
7 Keble Road, Oxford, OX1 3QL

BA PREVIEWS PROJECT

INSTRUCTION SHEET

REFERENCES FOUND BY SEARCHING OF CURRENT LITERATURE

A. GENERAL INSTRUCTIONS

1. Please use BLOCK LETTERS when completing the form.

2. Please list up to 30 references, found during the next six weeks by your normal methods of searching the current literature, which are relevant (MAJOR or MINOR value) to the research interests your profile is intended to serve.

3. The forms should be returned on or before the date given at the top of the form.

4. References may be obtained from any source except Biological Abstracts or your computer printout. Possible sources will include abstract journals (e.g. Microbiology Abstracts), titles lists (e.g. Current Contents), original (primary) journals or personal communication from a librarian or colleague.

   N.B. References found in an abstract journal, other than BA, are of particular interest for the experiment. If you use an abstract journal for current awareness purposes, please make a special effort to scan some of the current issues during the next six weeks and at least one issue if the abstract journal is published monthly or quarterly.

5. We realise that some participants will find considerably less than the 30 references which is the maximum number for which space is allowed on the form. However, if you expect to find more than 30 references during the time allowed, give preference to:

   (i) references located in abstract rather than primary journals;
   (ii) references located in earlier rather than later issues of primary journals;
   (iii) MAJOR rather than MINOR references.

B. INFORMATION REQUIRED ON THE FORM

6. For references to journal articles give first author, journal abbreviation, volume number (or year), issue number (if any) and first page number.

7. References to conference proceedings, books, reports and other types of 'non-journal' literature should be specified by the first author and sufficient additional information to enable us to identify the reference.

8. Source: Indicate where you located the reference, e.g. original (primary) journal, Chemical Abstracts, Forestry Abstracts, a library bulletin, personal communication.

9. Read: You are not being asked to read a paper before deciding its relevance. However, if you have read it, please put an 'X' in this column.

May 1971

BV-6
(Insts.)

xperimental Information Unit
/ Keble Road, Oxford

Sheet

BA-PREVIEWS PROJECT

REFERENCES FOUND BY SEARCHING CURRENT LITERATURE

rofile No: ....Name: .... Date to be returned: ....

1st author	Journal abbreviation	Vol. no.	Issue no.	1st page no.	Major or Minor	Source	Read (X)	Office use only			
Example: POLLARD	J MOL BIOL	50	12	91	MAJ	Current Cont	X				

BV-6

BA-PREVIEWS PROJECT

It is most important that the attached forms are completed and returned by the specified date if at all possible. However, some participants may not be able to do literature searching at any time during the next six weeks owing to holidays, field work or for other reasons. If delay is unavoidable, please keep the form (BV-6) but return this card indicating the month during which you expect to be able to undertake the searching.

Name: .... Profile No: ....

Month when searching could be undertaken ....

Experimental Information Unit
7 Keble Road, Oxford

BA-PREVIEWS PROJECT

CHEMICAL ABSTRACTS CONDENSATES (CAC) PRINTOUT

INSTRUCTION SHEET

GENERAL

1. The next stage of our evaluation involves a comparison of the coverage of BA-Previews with MEDLARS (the magnetic tape version of Index Medicus) and Chemical Abstracts Condensates (the tape version of Chemical Abstracts) for those evaluees whose subject interests are in the biomedical or biochemical field respectively.

2. For your own subject requirements, Chemical Abstracts Condensates (CAC) would appear to be the appropriate alternative service. We have therefore constructed a CAC profile to cover your research interests and this has been run against 4 weekly issues of CAC tapes.

3. Two copies of the CAC printout are enclosed. One copy is for your own use and the second copy for you to indicate the relevance of each reference retrieved.

INFORMATION REQUIRED

4. Will you please mark each reference as:-

MAJ (Major Value) / MIN (Minor Value) / X (Irrelevant)

5. Each reference should be marked in the outer margin, i.e. the lefthand column of references should be marked in the lefthand margin, the righthand column in the righthand margin.

6. We do not require you to indicate whether or not you already knew of the existence of an article.

7. The output should be returned in the enclosed addressed envelope, as soon as possible.

GROUP PROFILES

8. If more than one person is supplying relevance decisions, please follow exactly the procedure outlined in our previous instruction sheet for 'Relevant References Retrieved by the BA-Previews Service' (BV-7).

Experimental Information Unit
7 Keble Road, Oxford

BA-PREVIEWS PROJECT

MEDLARS PRINTOUT: INSTRUCTION SHEET

GENERAL

1. The next stage of our evaluation involves a comparison of the coverage of BA-Previews with MEDLARS (the magnetic tape version of Index Medicus) and Chemical Abstracts Condensates (the tape version of Chemical Abstracts) for those evaluees whose subject interests are in the biomedical or biochemical field, respectively.

2. For your own subject requirements, MEDLARS would appear to be the appropriate alternative service. We have therefore constructed a MEDLARS profile to cover your research interests and this has been run against a 3-month MEDLARS file.

3. Two copies of the MEDLARS printout are enclosed. One copy is for your own use and the second copy for you to indicate the relevance of each reference retrieved.

INFORMATION REQUIRED

4. Will you please mark each reference as:-

MAJ (Major Value) / MIN (Minor Value) / X (Irrelevant)

5. Each reference should be marked in the outer margin, i.e. the lefthand column of references should be marked in the lefthand margin, the righthand column in the righthand margin.

6. We do not require you to indicate whether or not you already knew of the existence of an article.

7. The output should be returned in the enclosed addressed envelope, as soon as possible.

GROUP PROFILES

8. If more than one person is supplying relevance decisions, please follow exactly the procedure outlined in our previous instruction sheet for 'Relevant References Retrieved by the BA-Previews Service' (BV-7).

Experimental Information Unit
7 Keble Road, Oxford OX1 3QL

ASCA SERVICE : INSTRUCTION SHEET

GENERAL

1. The next stage of our evaluation will involve a comparison of BA-Previews with other computerised information services. For those evaluees whose subject interests are in the biomedical field, we have constructed a MEDLARS profile and for those in the biochemical field a Chemical Abstracts Condensates profile. As your profile was not considered suitable for either of these, we would like to give you a citation search, using ASCA (Automatic Subject Citation Alert), the magnetic tape version of Science Citation Index.

2. The unusual nature of the ASCA search should result in the retrieval of a substantial number of articles which would not be found by a search of titles, or of conventional indexes, of the type used in the BA-Previews search.

3. The ASCA service allows a search to be made for cited authors or cited references.

   Cited author search: For a given author, this will retrieve all current articles which cite in their bibliographies any paper by this particular author. (It will not retrieve papers written by this author).

   Cited reference search: This will retrieve only current articles which cite this particular reference in their bibliographies.

INFORMATION REQUIRED

4. Will you please list on the attached form not less than 10, and not more than 15, cited authors and/or cited references.

   a) Cited author: a key author who in the past has published work relevant to your research work, and whose papers are likely to be cited in current articles which are relevant to you.

INSTRUCTION SHEET (CONT'D)

b) Cited reference: a specific reference by a key author. This should be given if any of your key authors has published in several research areas and only one or two of his papers are likely to be cited in papers relevant to your interests. Books (or specific chapters by individual authors) and other non-serial publications may be given, but we would prefer that you restrict this type to no more than 3 references.

N.B. Your list of citations may consist of authors only, individual references only, or a mixture of the two.

5. The list of citations should cover all the research interests included in your BA-Previews profile.

6. References published in 1970 or 1971 should not be included. (It is rare for references to be cited the year immediately following publication).

7. There is no limit to the age of cited references. However, as a general rule, references more than 10 years old are not often cit with the exception of very important authors or papers. (This may not be true of some subject areas, e.g. classical taxonomy.)

8. Will you please return the completed form in the enclosed addressed envelope, as soon as possible.

Experimental Information Unit
7 Keble Road Oxford OX1 3QL

BA-PREVIEWS PROJECT

ASCA SERVICE : CITATION SEARCH

INSTRUCTION SHEET

GENERAL

1. A few weeks ago you provided us with a list of up to 15 cited authors and/or references for a citation search.

2. We had hoped to give you a computer printout of references using ASCA, the magnetic tape version of Science Citation Index (S.C.I.). However, this proved impossible.

3. A manual search for the citations has therefore been made in the S.C.I. for January-March, 1971.

4. On each of the enclosed cards is a reference which has cited one of the authors or references you gave us.

INFORMATION REQUIRED

5. Will you please mark each reference:-

MAJ (Major Value) / MIN (Minor Value) / X (Irrelevant)

6. Each reference should be marked in the right-hand margin.

7. We do not require you to indicate whether or not you already know of the existence of the article.

8. The cards should then be returned in the enclosed envelope, as soon as possible.

9. If more than one person is supplying relevance decisions, please follow exactly the procedure outlined in our previous instruction sheet for "Relevant references retrieved by the BA-Previews Service". (BV-7)

Appendix IV.5

PROCEDURE FOR THE ANALYSIS OF COVERAGE SAMPLES

1. A reference was classified as covered only if the enriched title and full bibliography had been located in either the abstracts of BA or the bibliographic listing in BRI.

2. Assistance in locating references was provided by a pagination list which was generated by a computer search of the individual BA and BRI tapes for the period January 1971 to February 1972 inclusive. This list was in journal coden, then volume number, then issue number, and then page number order. It provided a record of the page range of each journal which appeared in each issue of BA or BRI. Thus, a sample entry is:

```
LYMPB   LYMPHOLOGY

      5220   2   4  160   166

      5205   3   2   71    75
```

   This entry shows that issue 20 of BA (vol 52) contained entries from the journal, Lymphology, volume 2, issue no. 4. The earliest page number to appear in the issue of BA was page 160 and the latest was 166.

3. The pagination list provided an indication of the probable issue of BA or BRI in which a particular reference in the coverage sample appeared. However, this list was not inerrant, nor did we assume it to be so. Also, we believe that a small but significant proportion of items from the coverage sample might appear in BA or BRI either before or after the 14 month period covered by the pagination file.

4. Items from journals which had contributed no items to the pagination list were checked against the BIOSIS source list for 1970-72 to determine whether the particular journal was covered at all by BA or BRI. Items from journals which were not covered were coded accordingly.

5. For all other items which had not been identified as being covered by either BA or BRI, a search was made by means of the cumulated author indexes of BA and BRI for 1970 and 1971. In the case of all the alternative service samples and about one half of the BV-6 samples author indexes of the individual issues of BA and BRI were searched for the period January 1972 - September 1972 inclusive. Articles from the remaining BV-6 samples were checked against the 1972 author indexes for January and February only.

6. In checking the author indexes as much care as possible was taken to avoid errors arising from alternative transliterations (e.g. from the cyrillic

alphabet).

7. Some references were also checked by the use of the Key-Word-In-Context (KWIC) indexes of BA and BRI.

8. Inevitably human errors will have occurred in using the author indexes and the pagination list. Furthermore some items which were classified as not covered may in fact have been covered outside the period which was searched. Nevertheless, we are satisfied (cf Appendix IV.6) that the great majority of items occurred during the period searched and, more particularly, in the 1971 issue of BA or BRI. It will be noted that these issues were checked by two largely independent methods - via the pagination list, and by means of the author index. We are therefore satisfied that the error rate in identifying covered references is unlikely to have a significant effect on our estimates of coverage ratios.

Appendix IV 6

TABLE: DISTRIBUTION OF APPEARANCE DATES IN BA PREVIEWS OF REFERENCES FROM COVERAGE SAMPLES

MONTH	SOURCE OF		SAMPLE	MEDLARS
	BV-6	SCI	CAC	
1970				
Jan-March	1	0	0	1
April-June	4	0	2	8
July-Sept	4	0	0	3
Oct-Dec	3	0	2	2
1971				
Jan	11	0	1	17
Feb	5	2	29	18
March	8	1	47	32
April	9	12	75	32
May	9	28	38	75
June	28	94	23	68
July	42	79	23	75
Aug	42	44	8	38
Sept	83	16	4	18
Oct	83	6	10	23
Nov	79	2	7	17
Dec	43	3	6	20
1972				
Jan	24	3	0	8
Feb	9	2	2	8
March	0	0	1	3
April	2	0	1	7
May	2	0	0	4
June	2	0	1	0
July	0	0	0	2
Aug	0	0	0	1
Sept	0	0	0	2
Total	493	292	280	482

The above table shows the chronological distribution of references in each of the four coverage samples which were covered by either Biological Abstracts or BioResearch Index. For each sample, the number of references appearing in each month of BA or BRI is given.

Data for the ASCA, CAC and MEDLARS samples are based on all references in the coverage samples which were covered by BA or BRI. The data for the BV-6 sample ('references found by the user') for a subset of 493 of the 950 references covered by the data base.

The references in the alternative service samples corresponded to the following issues of these services:

MEDLARS: April-June, 1971

CAC: Vol 74; issues 5, 7, 9, 11 (published in February/March, 1971)

SCI: January-March, 1971

BV-6 samples consisted of articles found by the user in June, July and August.

Appendix IV 7

Currency spread in BA/BRI for individual issues of primary journals

A random sample of 250 individual issues of primary journals was selected. For each issue, we identified the number of months of BA or BRI over which articles from that issue were distributed, i.e. a 'spread' of 5 months indicates that there was a gap of 5 months between the earliest and latest dates on which articles from a given journal issue appeared in BA/BRI.

The results are shown in the following table. For example the table shows that an individual journal issue was spread over 3 months of BA/BRI in 38 of the 250 issues in the sample.

TABLE: CURRENCY SPREAD IN BA/BRI

Currency Spread (Months)	No. of Occurrences	% of Occurrences
1	46	18.4
2	54	21.6
3	38	15.2
4	37	14.8
5	6	2.4
6	2	0.8
7	12	4.8
8	12	4.8
9	17	6.8
10	16	6.4
11	9	3.6
12	1	0.4
	250	100.0

Appendix IV.8

TABLE: SERIAL PUBLICATIONS NOT COVERED BY BIOLOGICAL ABSTRACTS OR BIORESEARCH INDEX

The following table gives the 86 journals which contributed one or more relevant items to the coverage samples but which were not covered by BIOSIS in 1971.

A journal coden is given for those titles which could be located in the ASTM listing. Journals for which no coden could be located are indicated as follows:

* the journal title was identified in one of the standard serials listings (Union List of Serials, Current Serials Received by the NLL, etc) and there is no doubt as to its identity.

? (11 titles) the title could not be found in any serials list. There is a possibility that either the title is incorrect or that it does not refer to a serial publication.

Journals contributed relevant items to one user only with the exceptions indicated by the figure in parenthesis after the coden, e.g. AVPZA(2): this journal contributed relevant items to coverage samples from two users.

Journal Title	Coden
Acta Parasitologica Lithuanica	APLUA
Aliment et la Vie	ALLVA
American Corrective Therapy Journal	*
Archiv fuer Pflanzenschutz	AVPZA (2)
Arkady Review	AKRVA
Atti della Societa Italiana Buiatria	*
Biokhim Immuniteta Pokoya Rast	?
Brauwelt	BRUWA
Bulletin of the American Ceramic Society	BACSA
Bull Fac Domestic Sci, Otsuma University	?
Bull Jap Ent Acad	?
Bulletin of the Japan Society of Mechanical Engineers	BJSEA
Bulletin de Madagascar	BUMDA
Bulletin of Prosthetics Research	BPRRB
Cellular Immunology	*
Ceskoslovenska Neurologie	CSKNA
Ceskoslovenska Zdravotnictvi	CEZDA
Chemical Technology, A Series of Monographs	CHTNA (3)
Chemist and Druggist	CHDRA
Current Topics in Pathology	*
Deutsche Zeitschrift fur Verdauungs und Stoffwechselkrankheiten	DZVSA
Diagn Potrebnosti Rust Udobr	?
Doklady Kazakhskoi Akademii Sel/Skokhozyaistvennykh Nauk	DKASA
Doklady Bulgarska Akademii na Naukite	DBANA
Essays Toxicol	?
Fishery Bulletin of the National Oceanic and Atmospheric Administration	*
Fiziol Aktiv Komponenty Pitan Zhivotn	?
Fiziologiya i Biokhimiya Rastenii Akademii nauk Belorusskoi SSR	FBIBA
Fleischwirtshaft	FLEIA
Food World	FOWDA

Journal Title	Coden
Gen Eng (India)	?
Genetika a Slechteni	GESLB
Geochemistry International	GCINA
Geochimica et Cosmochimica Acta	GCACA
Geotechnica	GETCA
I.S.I. Bulletin (Indian Standards Institution)	ISTBA
Industria Alimentara	INALA
Industries Alimentaires et Agricoles	IALAA
Inland Acad Sci (Indiana)	*
Int Journal of Parasit	?
Izvestiya na Mikrobiologicheskiya Institut...	IMNSA
Japan Agricultural Research Quarterly	*
Japanese Society of Internal Medicine - Journal	AJIMA
Journal of the Agricultural Association of China	JAGAA
Journal of the Institution of Chemistry	JOICA
Journal of the Society of Dairy Technology	JSDTA
Journal of Sports Medicine and Physical Fitness	JMPFA
Kagaku No Ryoiki Zokan	KNRZA
Kharchova Promislovist, Naukovo-Tekhnichnii Zbirnik	KPTZA
Kvasny Prumysl	KVPRA
Lambillionea	LMBLA
Minerva Ginecologica	MIGIA
Minerva Nipiologica	MINIA
Nederlandsch Tijdschrift voor Verloskunde en Gynaecologie	NTVGA (2)
Nerv Gumoral Mekh Regul Funkts	?
Nippon Dojo-Hiryogaku Zasshi	NIDHA
Nippon Sakumotsu Gakkai Kiji	NISAA
Pesticide Science	*
Poljoprivredna Znanstvena Smotra	PJZSA
Problemy Parazitologii	PPUZA
Progress in Phytochemistry	PRPCB
Regul Rosta Rast Khim Sredstvami	?
Revista Chilena de Obstetricia y Ginecologia	RCOBA
Ricerca Scientifica, Revista	RSRVA (2)
Royal Society of Health Journal	RSHEA (3)

Journal Title	Coden
Scandinavian Journal of Dental Research	*
Sedimentary Geology	SEGEB
Seibutsu-Butsuri	SEBUA
Studia Biophysica	STBIB
Studii si Cercetari de Medicina Interna	SCMDA
Suplemento Cientifico. Consejo General de Colegios Veterinarios de Espana	*
Teor Vop Obrab Pochv	?
Tombo	*
Triangle	TRGLA
Tropical Diseases Bulletin	TDBUA
Vestnik Moskovskogo Universiteta Seriya II Khimiya	VMKUA
Vestnik Studencheskogo Nauchnogo Obshchestva, Kazanskii Gosudarstvennyi Universitet Im V.I. Ul/Yanova-Lenina	VOKUA
Vestsi Akademii Navuk Belaruskai SSR, Seriya Biyalagichniykh Navuk	VABBA (3)
Veteriner Fakultesi Dergisi	VTFDA
Voprosy Fiziologii i Biokhimii Kul/Turnykh Rastenii Akademiya Nauk Moldavskoi SSR, Institut Fiziologii i Biokhimii Rastenii	VFBRA
Waldhygiene	WLDHA
Zeitschrift fuer Geburtshilfe und Gynaekologie	ZGGNA (2)
Zesyty Naukowe Uniwersytetu W Toruniu, Nauki Matematyczno-Przyrodnicze-Biologia	ZTMBA
Zhurnal Analiticheskoi Khimii	ZAKHA
Zhurnal Organicheskoi Khimii	ZORKA
Zivotnovadni Nauki	*

Appendix V

CONTENTS

1. FINAL QUESTIONNAIRE
2. ADDITIONAL ADVANTAGES AND DISADVANTAGES OF THE BA PREVIEWS SERVICE WHICH WERE MENTIONED BY USERS IN REPLY TO QUESTION 1
3. TABLE: CORRELATIONS BETWEEN USERS' REPLIES TO QUESTIONS 12 AND 13 AND REPLIES TO QUESTIONS 1-11
4. CARD OUTPUT: [QUESTIONNAIRE]

Experimental Information Unit
7 Keble Road, Oxford

BA-PREVIEWS PROJECT

FINAL QUESTIONNAIRE

BEST COPY AVAILABLE

Profile No: .... Name: ....

1. We should like to find out what you consider to be the main advantage and disadvantages of the BA-Previews service. For this purpose we have listed below some of the possibilities. Will you please tick only those which you consider to be really IMPORTANT to you. Do not tick any which are only of minor significance. You may tick none, one or several in each list. Others may be added in the spaces provided.

ADVANTAGES

a) The profile retrieves references of major value which I might not otherwise locate. ☐

b) The profile retrieves background references which I might not otherwise locate. ☐

c) The service covers journals which I would not normally see. ☐

d) The service retrieves references from non-journal publications, e.g. books, symposia, conference proceedings,etc. ☐

e) The service is time-saving, i.e. it reduces the time spent searching the literature by normal methods. ☐

f) The service is convenient, i.e. it obviates the need for copying out references from abstracting and indexing journals for incorporation in my personal index. ☐

g) The provision of a regular service gives me confidence; my literature searching is no longer dependent upon how and when I can do it myself. ☐

Others (please specify):-

h) ....

.... ☐

....

i) ....

.... ☐

....

DISADVANTAGES

j) The profile retrieves too many irrelevant references ☐

k) The profile misses too many relevant references ☐

l) There is too much delay between publication of an article and its appearance in BA-Previews ☐

m) The journal coverage is inadequate ☐

n) Many relevant references are of little or no practical value:

(i) They are from journals which are not easily accessible ☐

(ii) They are from publications which I consider to be of low quality ☐

(iii) They are from foreign language journals which I cannot read or afford to have translated ☐

o) The computer printout is unattractive ☐

p) The computer printout is difficult to file ☐

q) The coding of the profile is difficult to understand so I am not sure how the profile logic is intended to work, i.e. which single search term or combination of search terms will retrieve a reference ☐

r) There is not enough information on the printout:

(i) no abstracts ☐

(ii) no author addresses ☐

(iii) no indication of the language of the original paper ☐

Others (please specify):-

s) ....

.... ☐

....

t) ....

.... ☐

....

/cont...

BA-Previews Project: Questionnaire

Note: Questions 2, 3 and 4 (below) refer to disadvantages listed in Question 1. However, we would like a direct answer, irrespective of whether you have ticked them as 'Important Disadvantages'.

2. Do you consider that your profile retrieves too many irrelevant references?
   - Yes ☐
   - No ☐
   - Don't know ☐

3. Do you consider that your profile misses too many relevant references?
   - Yes ☐
   - No ☐
   - Don't know ☐

4. Do you consider the journal coverage of BA-Previews for your research interests is:
   - Adequate ☐
   - Inadequate ☐
   - Don't know ☐

5. Do you find it difficult to decide, from the title alone, whether an article is relevant to your research interests?
   - Frequently ☐
   - Occasionally ☐
   - No ☐
   - Don't know ☐

6. The use of a liaison scientist to provide assistance in the construction and amendment of profiles is expensive. Furthermore, some research workers may not wish to delegate this task to someone else but would prefer to construct their own profile. Would you prefer:-
   - a) To construct your own profile, after receiving the necessary instruction on the technique of profile construction, and have it checked, or ☐
   - b) To have the profile constructed by a liaison scientist, as at present. ☐

/cont...

7. Since you started receiving BA-Previews computer printouts, has there been an increase or decrease in your use of the printed Biological Abstracts for Current Awareness and/or Retrospective searching?

	Current Awareness	Retrospective
Increase		
Decrease		
No change		
Don't know		
Never use it		

8. After selecting the relevant references on your computer printout, do you subsequently look up the original articles, provided you have not already seen them?

Always ☐
Frequently ☐
Occasionally ☐
Never ☐

9. Have you experienced difficulty in obtaining the original articles?

Always ☐
Frequently ☐
Occasionally ☐
Never ☐

10. Some participants have asked if there will be a permanent BA-Previews Current Awareness service. If there were such a service do you consider that you would be able to reduce the time spent:-

a) Scanning current journals

Considerably ☐
Slightly ☐
No ☐
Don't know ☐

b) Scanning secondary journals, i.e. abstracting and indexing journals and titles lists

Considerably ☐
Slightly ☐
No ☐
Don't know ☐

11.	If, in addition to the Current Awareness service it became possible to operate computer searches of Biological Abstracts for the last 1-5 years, would this Retrospective search service be of greater or lesser value to you than the existing Current Awareness service?	Greater value The same value Lesser value Don't know
12.	If there were a permanent BA-Previews Current Awareness service would you wish to become a subscriber, assuming that your department/ institution would pay the subscription?	Yes No Don't know
13.	If a permanent BA-Previews Current Awareness service were established in the UK, on a full cost-recovery basis, the minimum cost of a profile is likely to be £50-£75 per annum (though improvements in search programs might reduce this substantially during the next few years). Do you consider your present BA-Previews profile to be worth £50-£75 per annum?	Yes No Don't know
14.	How many people (including yourself) see your printout regularly?	....

FURTHER COMMENTS:-

Appendix V 2

Additional advantages and disadvantages of the BA Previews service which were mentioned by users in reply to Question 1.

In the final questionnaire users were provided with lists of both advantages and disadvantages and were asked to tick those which they considered to be applicable. They were also given the opportunity to nominate advantages and disadvantages other than those listed. Only three advantages and three disadvantages were specified in the 'others' category and these are listed below.

Advantages

1. "In comparison with UKCIS [CA-condensates service], the use of enrichment terms enables me to identify more efficiently articles of major interest."
2. "The service brings to my notice rapidly developing techniques, materials, etc. in fields other than those in which I work and which may be useful in my own field."
3. "The service acts as a recurring stimulus to look at current publications more frequently than I would otherwise."

Disadvantages

1. "Inaccuracies in the printout, such as misquotations of authors names and confusion of taxonomic terms, especially similar generic names, are frequent and reduce confidence in the system."
2. "There are poor translations of foreign-language titles, especially of technical terms."
3. "It is not always possible to tell in which country the work was done."

Appendix V 3

TABLE: CORRELATIONS BETWEEN USERS' REPLIES TO QUESTIONS 12-13 AND REPLIES TO QUESTIONS 1-11

QUESTION	QUESTION 12			QUESTION 13		
	YES %	NO %	α	YES %	NO %	α
Q1 **Advantages**						
1(a) YES	60	21	0.0001	62	39	0.0001
1(b) YES	76	58	0.003	74	68	(0.34)
1(c) YES	81	72	(0.13)	78	78	(0.97)
1(d) YES	39	33	(0.40)	39	35	(0.63)
1(e) YES	47	20	0.0001	54	29	0.0001
1(f) YES	22	1	0.0001	25	10	0.001
1(g) YES	40	12	0.0001	47	21	0.0001
Q1 **Disadvantages**						
1(j) YES	39	52	0.05	32	50	0.003
1(k) YES	9	12	(0.60)	5	13	(0.07)
1(l) YES	45	52	(0.30)	39	53	0.03
1(m) YES	8	12	(0.50)	9	10	(1.0)
1(n - i) YES	16	20	(0.55)	12	21	0.06
1(n - ii) YES	9	7	(0.69)	8	9	(0.94)
1(n - iii) YES	22	21	(0.98)	15	26	0.05
1(o) YES	29	26	(0.67)	32	25	(0.27)
1(p) YES	42	34	(0.26)	43	37	(0.44)
1(q) YES	11	10	(0.90)	11	11	(0.88)
1(r - i) YES	29	21	(0.18)	26	28	(0.83)
1(r - ii) YES	30	21	(0.16)	27	28	(0.96)
1(r - iii) YES	40	30	(0.17)	35	38	(0.79)

BEST COPY AVAILABLE

TABLE (CONT'D)

QUESTION		QUESTION 12			QUESTION 13		
		YES %	NO %	α	YES %	NO %	α
Q2	YES	46	64	0.01	39	60	0.001
Q3	YES	13	14	(0.96)	8	17	0.05
Q4	ADEQUATE	74	69	(0.41)	76	70	(0.33)
Q5 -	FREQUENTLY	13	20	(0.19)	12	17	(0.24)
	FREQUENTLY/ OCCASIONALLY	80	83	(0.72)	79	82	(0.60)
Q8	ALWAYS	16	20	(0.55)	17	17	(0.97)
	ALWAYS/FREQUENTLY	71	63	(0.21)	73	66	(0.29)
Q10 (a) or (b)	CONSIDERABLY	70	33	0.0001	73	49	0.0001
Q11	GREATLY				35	47	(0.06)

Notes to Table:

(i) the question numbers listed in the left-hand column are these used in the final questionnaire (See Appendix V.1)

(ii) the percentages indicate the proportion of those users giving a particular reply to question 12 or 13 who also gave the reply indicated in the first column. E.g. the first entry in the column headed 'QUESTION 12 - YES' indicates that 60% of users replying 'Yes' to question 12 had also ticked advantage 1(a)

(iii) the significance α was computed from values of χ^2. A correlation was considered to be 'not significant' if $\alpha > 0.05$ and the value of α is shown in parentheses (two-tailed test)

Appendix V 4

Experimental Information Unit,
7 Keble Road, Oxford OX1 3QL

BA-PREVIEWS : CARD OUTPUT

During the course of the BA-Previews experiment, some considerable interest has been shown by participants in the possibility of the computer printout being available on cards as opposed to paper.

In order to demonstrate the alternative format, this issue of BA-Previews has been printed on cards. If the BA-Previews service continues for a further experimental period, it is likely that there will be a choice between card or paper output.

We should like to have your opinion of this card output and would be grateful if you would kindly tick the appropriate box below and return this form to the above address.

1. I prefer the card output ☐
2. I prefer the paper output ☐
3. I have no preference for cards or paper ☐

Appendix VI

CONTENTS

1. INFORMATION CENTRES: FINAL QUESTIONNAIRE

Appendix VI 1

Experimental Information Unit
7 Keble Road

☐☐☐

BA-PREVIEWS PROJECT

INFORMATION CENTRES

FINAL QUESTIONNAIRE

Profile No: .... Name: ....

1. We should like to find out what you consider to be the main advantages and disadvantages of the BA-Previews service. For this purpose we have listed below some of the possibilitiss. Will you please tick only those which you consider to be really IMPORTANT to you. Do not tick any which are only of minor significance. You may tick none, one or several in each list. Others may be added in the spaces provided.

ADVANTAGES

a) The profile retrieves references which we might not otherwise locate. ☐

b) The profile retrieves articles from journals which we do not cover in our existing literature searching. ☐

c) The service retrieves references from non-journal publications, e.g. books, symposia, conference proceedings, etc. ☐

d) The service is time-saving, i.e. it obviates the need for manual searching of the hard copy Biological Abstracts. ☐

e) The service is convenient, i.e. it obviates the need for copying out references from secondary journals, as the print-out can be used for processing our bulletin/index. ☐

Others (please specify):-

f) ....

....

.... ☐

g) ....

....

.... ☐

g) The profile retrieves too many irrelevant references. ☐

h) The profile misses too many relevant references. ☐

i) There is too much delay between publications of an article and its appearance in BA-Previews. ☐

j) The journal coverage is inadequate. ☐

k) The computer printout is unattractive. ☐

l) The computer printout is difficult to file. ☐

m) The coding of the profile is difficult to understand. ☐

n) There is not enough information on the printout: ☐
  - (i) no abstracts ☐
  - (ii) no author addresses ☐
  - (iii) no indication of the language of the original paper. ☐

Others (please specify):-

o) ....

....

.... ☐

p) ....

....

.... ☐

/cont....

BA-Previews Project: Questionnaire

Note: Questions 2, 3 and 4 (below) refer to disadvantages listed in Question 1. However, we would like a direct answer, irrespective of whether you have ticked them as 'Important Disadvantages'.

2. Do you consider that your profile retrieves too many irrelevant references?

Yes ☐

No ☐

Don't know ☐

3. Do you consider that your profile misses too many relevant references?

Yes ☐

No ☐

Don't know ☐

4. Do you consider the journal coverage of BA-Previews for your Information Centre's interest is:

Adequate ☐

Inadequate ☐

Don't know ☐

5. Do you find it difficult to decide, from the title alone, whether an article is relevant to the interests of your Information Centre?

Frequently ☐

Occasionally ☐

No ☐

Don't know ☐

6. The use of a liaison scientist to provide assistance in the construction and amendment of profiles is expensive. Furthermore, you may not wish to delegate this task to someone else but would prefer to construct your own profile. Would you prefer:-

a) To construct your own profile, after receiving the necessary instruction on the technique of profile construction, and have it checked, or ☐

b) To have the profile constructed by a liaison scientist, as at present. ☐

/cont....

7. Since you started receiving BA-Previews computer printouts, has there been an increase or decrease in your use of the printed Biological Abstracts for Current Awareness and/or Retrospective searching?

	Current Awareness	Retrospective
Increase		
Decrease		
No change		
Don't know		
Never use it		

8. Some participants have asked if there will be a permanent BA-Previews Current Awareness service. If there were such a service do you consider that you would be able to reduce the time spent by your Information Centre:

a) Scanning current journals

Considerably ☐

Slightly ☐

No ☐

Don't know ☐

b) Scanning secondary journals

Considerably ☐

Slightly ☐

No ☐

Don't know ☐

9. If, in addition to the Current Awareness service it became possible to operate computer searches of Biological Abstracts for the last 1-5 years, would this Retrospective search service be of greater or lesser value to you than the existing Current Awareness service?

Greater value ☐

The same value ☐

Lesser value ☐

Don't know ☐

Cont/ ....

BEST COPY AVAILABLE

10. If there were a permanent BA-Previews Current Awareness service would you wish to become a subscriber, assuming that adequate funds were available?

Yes ☐

No ☐

Don't know ☐

11. If a permanent BA-Previews Current Awareness service were established in the UK, on a full cost-recovery basis, the minimum cost of a profile is likely to be £50-£75 per annum (though improvements in search programs might reduce this substantially during the next few years). Do you consider your present BA-Previews profile to be worth £50-£75 per annum?

Yes ☐

No ☐

Don't know ☐

12. We should like to know what use, if any, you have been making of your BA-Previews printout during the experiment. Some possible uses are listed below. Will you please tick those which apply to your Centre and also supply the missing details, as indicated. Others may be added in the spaces provided.

a) *All/Some of the relevant references are selected for incorporation in a published *Abstract Journal/Information Bulletin, which is distributed to .... subscribers.

b) *All/Some of the relevant references are selected for incorporation in an Information Bulletin, which is distributed free to .... workers.

c) *All/Some of the relevant references are selected for incorporation in the Information Centre's own card index.

* Delete whichever does not apply

d) None of the references on the printout have been used by the Information Centre.

Others (please specify):-

e) ....

....

....

f) ....

....

....

cont/ ....

FURTHER COMMENTS:-